The British Journal of Educational Psychology

(Incorporating the "Forum of Education")

EDITED BY

C. W. VALENTINE

WITH THE ASSISTANCE OF

R. L. ARCHER
P. B. BALLARD
SIR CYRIL BURT
JAMES DREVER
LL. WYNN JONES
H. BOMPAS SMITH
SIR GODFREY THOMSON
P. E. VERNON
HELEN M. WODEHOUSE

AND THE FOLLOWING REPRESENTATIVES OF OTHER COUNTRIES:

L. CARMICHAEL
ARNOLD GESELL
I. MEYERSON
PETER PETERSEN
JEAN PIAGET
FRED J. SCHONELL
L. M. TERMAN

Volume XX

1950

Published by
METHUEN AND CO., LTD., 36, Essex Street, Strand, London, W.C.2,
for the British Journal of Educational Psychology, Ltd.
Registered Offices: 42, Hill Street, Birmingham, 5.
on behalf of
THE BRITISH PSYCHOLOGICAL SOCIETY and THE ASSOCIATION
OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION.

VOLUME XX. FEBRUARY 1950 PART I.

SYMPOSIUM ON THE SELECTION OF PUPILS FOR DIFFERENT TYPES OF SECONDARY SCHOOLS.

IX.—CONCLUSION.

BY SIR CYRIL BURT
(University College, London).

I.—*Points of agreement*: (*i*) *The need for a comprehensive basis*; (*ii*) *the importance of innate abilities*; (*iii*) *the value of intelligence tests.* II.—*Points in doubt or dispute*: (*i*) *The composite nature of the practical factor*; (*ii*) *the contribution of general ability to performances in specific tests*; (*iii*) *the late emergence of the practical factors*; (*iv*) *the amount of differentiation at* 11 *plus*; (*v*) *no proof of marked differences*; (*vi*) *the study of practicable procedures*; (*vii*) *the quota principle.* III.—*Summary and conclusions.*

I.—POINTS OF AGREEMENT.

THE Editor has generously asked me, as the opener of this Symposium, to reply to criticisms and sum up the main conclusions to be drawn from the various papers that have appeared. May I first briefly formulate the points on which practically all the contributors seemed agreed, and then discuss in greater detail the points of apparent disagreement?

(i) *The Need for a Comprehensive Basis.*—To begin with, there is plainly a general consensus that the selection of pupils should be made, not simply on intellectual grounds alone, such as would be supplied by a scholastic examination and by tests of general intelligence and special aptitudes, but on as wide and comprehensive a basis as is practicable. This, I take it, is the gist of the first of Dr. Alexander's main conclusions: " The allocation should be based on tests of the child's abilities, aptitudes, and attainments, together with assessments of of the more important character factors and in the light of his cumulative record in school, with special reference to health factors; . . . sociological factors cannot be ignored." In regard to the sociological factors, he sounds a timely warning. There may be, he says, unsuspected dangers in placing too strong an emphasis on this consideration. Writing " from an administrator's point of view," he insists that " it is imperative that the procedure adopted should be acceptable to public opinion; and for that reason the greatest emphasis must fall on objective criteria." With that we should all agree. To borrow a celebrated legal dictum, it is not enough to do justice to each pupil; " justice must also *seem* to be done."

(ii) *The Importance of Innate Abilities.*—The second point which every contributor has either stressed or assumed is the importance of considering inborn ability. Once again I quote Dr. Alexander's own formulation. As regards intellectual grounds he says, " I agree with Burt that it is important to distinguish between the child's *acquired* attainments, as measured by examination in English, Arithmetic, and so on, and the child's *innate* abilities." The general opinion seems to be that the greatest weight should be attached to inborn capacities, and that, on the whole, it is wiser to use examinations of acquired attainment merely as qualifying criteria rather than as criteria for selection.

(iii) *The Value of Intelligence Tests.*—Thirdly, all the contributors have recognized that one of the most essential features in the process of selection should be the use of intelligence tests. This agreement deserves to be specially underlined, because one or two books recently published have suggested to the minds of many teachers and officials that there has been a radical change of view about the value of such tests.

Dr. Blackburn, for example, has declared that it has been " all too readily assumed that intelligence tests are measuring innate intelligence, and nothing else " ; whereas " all that has been shown is, in fact, that they measure the ability to answer intelligence test questions." Even at the lower end of the scale he thinks their use may be attended with considerable risk. Since, so he says, the psychologist's tests are commonly couched in verbal form, a man who is " an almost illiterate mechanical genius may find himself incarcerated in a home for mental defectives."[1] Similarly, Dr. Fleming points out that " recent researches " have shown how the subsequent progress of " many pupils " may fail to conform with the predictions based on intelligence tests applied five or ten years earlier. " In the first few decades of the twentieth century, Pearson, Burt, and other workers emphasized the importance of innate differences . . . and did not hesitate to assume that pupils who did well at an early age would maintain their relative superiority in later years " ; thus " twenty-five years ago, when evidence had not been obtained as to the individual variations which they concealed, it became customary to speak of the ' constancy ' of I.Q.s." Later studies, we are told, have demonstrated that " the degree of constancy revealed . . . is very much less than was anticipated in the nineteen twenties." This change of view, Dr. Fleming contends, is " of great importance in discussions of selection for secondary education ; and the educational implications of this have as yet scarcely been formulated."[2]

These and similar statements have been so often quoted in teachers' journals and at teachers' conferences that it is well to stress the fact that (as the reviews of the books quoted would sufficiently show) very few educational psychologists would accept such contentions. No psychologist has ever supposed that intelligence tests furnish *perfect* measurements of innate ability, any more than a clock or watch gives a perfect measure of the lapse of time. The point at issue is not—Are our tests perfect ? but—How large are their admitted imperfections ? Are they really large enough to invalidate their use for practical purposes ? It is quite wrong to suggest, as Dr. Blackburn does, that earlier writers have assumed that intelligence tests measure " *nothing else* " besides innate intelligence. At the end of his chapter he gives references to my own books and those of about a dozen other testers. But he will find no such assumption made in any one of them (though it has often been attributed to them in popular writings). It is equally erroneous to state, as Dr. Fleming does, that twenty-five years ago no evidence had been obtained on the inconstancy of the I.Q., and that in consequence early workers assumed that it was perfectly constant. Once again the question is—How great is the constancy or the lack of it, and is it too low to warrant the use of intelligence tests as *one* important item in deciding the allocation of pupils ? And, in the answers given, there has been no such radical change of outlook as Dr. Fleming and Dr. Blackburn imply.[3]

Looking back, then, at the views expressed by those who have contributed to this Symposium, we may, I think, conclude that the vast majority of psychologists competent to express an opinion would unhesitatingly accept Dr. Alexander's qualified conclusion, namely, that " at the age of 11+, innate general intelligence can be estimated with reasonable accuracy by means of standardized tests."

---

[1] BLACKBURN, J. : *Psychology and the Social Pattern*, 1945, pp. 8, 68.

[2] FLEMING, C. M. : *Adolescence*, 1947 ; pp. 108, 110-111.

[3] Since both these critics cite my own early writings, may I point out that I for one, even " in the nineteen twenties," secured a good deal of evidence as to individual variations in I.Q.s, and emphasized that the constancy was " but imperfectly realized " ? And I further contended that " tests, infinitely more scientific than those set out [in my book], can be but the beginning, never the end, of the examination of the child " (*Mental and Scholastic Tests*, 1920, pp. xv, 150f.). Thanks to more recent research, the tests that are available to-day may legitimately claim greater reliability and validity.

II.—POINTS IN DOUBT OR DISPUTE.

The chief question on which opinions differ most widely is the degree of importance attached to attempts at measuring special aptitudes, particularly aptitudes that might decide the allocation of the child to a grammar school rather than a technical school, or *vice versa.* It will perhaps be convenient to take the arguments advanced by the several contributions so far as possible in the order of their appearance.

(i) *The Composite Nature of the Practical Factor.*—Dr. Alexander observes that " there are two points in which I differ from Burt in this field." The first turns on the fact that " Burt seems to assume that the *F* factor and the *k* factor denote much the same composite factor." Alexander relies mainly on the work of Drew, which " seemed to establish that the *F* and *k* factors " (' the practical ' and ' spatial ' factors) " were independent, and confirmed the view that the *F* factor was capable of measurement at 11+, and that the *k* factor did not emerge until the age of 13+ or later." In a later contribution, however, Mr. Dempster maintained that in his own results " spatial tests and performance tests seem to measure the same ability." To my mind both conclusions, stated in this way, may be a little misleading. I myself do not believe that the factors indicated are quite the same, nor yet that they are quite independent. Such formulations are too simple and too sweeping. My own opinion is best expressed in terms of what I have called the theory of sub-divided factors. The results that my fellow workers and I have obtained suggest that there is first a fairly broad group factor—usually called a ' practical factor '—common to both types of activity, and secondly that this common factor tends to differentiate, during the later years of mental growth, into at least two sub-factors—a perceptual or ' visuo-spatial ' factor and a factor of ' mechanical ' aptitude. Of the two, the evidence would seem to indicate that the former commonly matures first ; the latter is more complex, and does not depend solely on maturation ; as ordinarily tested, it involves, for example, a factor for neuro-muscular skill, often largely acquired by experience. Indeed, I should hold that both are capable of further sub-division in a suitably planned research.[1]

(ii) *The Contribution of General Ability to Performance in Specific Tests.*—" The second point in this field," says Dr. Alexander, " on which I differ from Burt to some extent is on the correlation between tests of technical ability and tests of academic ability." I had assumed a correlation of 0·50; Dr. Alexander estimates it as only about 0·30. In other words, I hold that, at the age of eleven, the general factor plays a much larger part in work in specialized directions than he does. The lower figure would, of course, make it much easier to discover children with a marked academic or technical bias. But, unless the reliability of both sets of tests is very poor or the children tested do not form a genuinely random group, a figure of only 0·30 would seem to be decidedly exceptional.[2] However, as Dr. Alexander himself observes, " our difference is one of degree, and not of principle."

(iii) *The Late Emergence of the Practical Factor.*—In my first paper I maintained that " at the age of eleven, special aptitudes and interests, especially those of a practical, mechanical, or technical character, cannot be accurately assessed, except in a comparatively small proportion of cases." The majority of children show no marked degree of specialization until puberty or even later. Dr. Alexander believes I have overstated my case; and Professor Peel, in a fresh investigation, has endeavoured to discover whether a practical factor does emerge as early as eleven plus. His conclusion is that, "if a practical factor can be said to exist at the age of thirteen, it is equally evident at eleven."

---

[1] I need say no more on this here since I have summarized what I take to be the best established results in a recent issue of this *Journal* (XIX, pp. 187-190).

[2] So far as my own knowledge goes, in the few investigations reporting correlations as low as this, it has usually turned out that the investigator has not applied his tests to a complete or random age-group. Often he has tested only a single school, or a small group of schools in a single neighbourhood ; frequently he has tested only the children in certain school classes. In such cases, of course, the effect of selection will have been to reduce the correlation actually found well below the ' true ' value obtaining in the general population.

Actually, Professor Peel's figures do not show *equal* evidence. With his thirteen-year-old pupils his practical factor accounts for more of the variance than it does with those aged eleven or twelve. Admittedly the difference is slight. Yet it is to be noted that, with his eleven-year-olds, the contribution of general ability to the total variance happens to be *smaller* than at any later age. Now, in his introductory section, he himself has argued that "three types of research—factorial, clinical, and neurological"—all agree in indicating that the general ability will, as a rule, account for *more* of the variance at earlier ages than at later. Hence, his groups must be exceptional. Moreover, if I understand his factorial procedure correctly, what he describes as a "practical factor" is really a bipolar factor, contrasting practical ability with verbal ability. In passing he tells us that his youngest age-group differed from the 12- and 13-year-old group by containing "children who would later leave for the grammar school." The presence of these children, therefore, would surely tend to magnify the influence of the bipolar factor he obtains. Further, he has included in his analysis data from both boys and girls without making any allowance for sex. Now, even at eleven, boys show greater aptitude and interest in practical problems, while girls show a marked verbal superiority. Much of his bipolar factor may consequently be due to the sex difference. Finally, let me say that (and I imagine Dr. Peel would agree), I doubt whether any wide generalization of this kind can finally be based on results derived from a single school alone.[1]

(iv) *The Amount of Differentiation at Eleven Plus.*—One of the most instructive and original contributions is that made by Miss Constance Lambert, in her capacity as Examinations Officer to the Northumberland Education Authority. To me her observations were of special interest, because they seemed to throw light on some of the peculiarities I noted, when, many years ago, I was asked to carry out an examination in the same area by means of mental and scholastic group tests which I had previously used in London. The results we have both obtained would appear to suggest that in different educational areas the problem may present itself in somewhat different forms; hence it is not always safe to assume that data procured in one district will hold good for all others.

Miss Lambert emphasizes that it is important to consider, not only the pupil's apparent abilities, but also his interests. With this as a theoretical principle I should heartily concur. But I doubt whether we yet possess sufficient evidence to show how far the interests manifested at the age of 11 are likely to be permanent. This is plainly a matter on which research is urgently needed. Miss Lambert's suggestion of using a test of general information points to one line of approach. In certain fields, information, especially when of a kind acquired outside the school rather than within, provides a fairly trustworthy clue to the child's current or recent interests. In London, however, when we have carried out tests of general information on the same individuals at intervals of two or three years, we have often found well marked changes, particularly during the pre-pubertal stage. Indeed, one of Miss Lambert's own findings is that, with increasing age and experience, the bias of interest alters, veering more towards technical and away from academic subjects.

---

[1] In passing I should like to express my appreciation of the two novel and interesting tests that Professor Peel has described. These should undoubtedly prove a valuable addition to the somewhat limited collection of practical tests available for such work. It is, however, a little disappointing to find that their correlation with "school practical subjects" is so low (0·263).

On the other hand, I cannot agree that the method of factor analysis that he has adopted was the most appropriate for his purpose. He has apparently used Thurstone's 'centroid method,' but without Thurstone's supplementary rotation. However, Thurstone's mode of dealing with the communalities and sign-reversal is apt at times to introduce quite arbitrary changes in the factor variance (with which that method was not primarily concerned). I should agree that, if Professor Peel's purpose had been to examine the differentiation between a verbal and a practical bias at different ages, bipolar factors would have been adequate. But, as he states his problem, his chief concern was with the maturation of the practical factor itself; and, since the practical factor is a group factor, a group factor method would obviously seem to have been the more suitable procedure.

In her main investigation, along with her tests of general intelligence, she has included tests of the type used by Professor Peel. The main conclusion drawn is " the existence of the bipolarity postulated in the Norwood Report." She believes that there is a group of children ' with grammar school ability and a significant bias towards verbal ability and academic interest,' and another group of children ' of technical school ability, with a significant bias towards practical ability and technical interest.' However, no one, I think, would deny that in a reasonably large and unselected sample some bipolarity of this nature can be discerned, even at the age of 11. What is in question is its relative size and importance during these earlier years. If I have understood her figures correctly, Miss Lambert would be inclined to guess that about 30 per cent. of those who tested fall definitely into the former group, and another 30 per cent. into the latter group. But those she has tested themselves form a highly selected batch, namely, " the top 30 per cent. of the population, as determined by an examination in intelligence, English, and Arithmetic. Hence in the general population the proportions would be comparatively small.

The patterns of factor saturations that she prints as having been obtained with Thurstone's method of analysis would suggest that three groups are also present in the tests employed. However, such a suggestion does not seem borne out by a rough inspection of the pattern of coefficients in her correlation table. But the coefficients themselves seem somewhat attenuated by the effects of selection ; and I venture to think that a group-factor analysis, with tentative corrections for selection, might furnish a clearer result.

Let me say at once, however, that I do not think that factorial analyses are likely to give the final answer to our problems. Such analyses are necessary, or at any rate useful, because they indicate the nature of the types or tendencies to be discerned among the pupils at the age investigated ; and they reveal the nature and the relative importance of the general and specific abilities entering into the various tests. But to my mind factorial studies should be supplemented by multiple regressions with external criteria. This older and alternative approach is still too much neglected. Yet, for practical as distinct from theoretical purposes, it often yields a far more direct solution. In this connection it may be instructive to recall the experience gained in the analogous problem of personnel selection among recruits for the Army. There, too, I ventured to suggest[1] that a multiple regression technique would be likely to give a somewhat more direct answer to the practical issues involved ; and the results progressively accumulated have fully confirmed the value of this procedure.

Miss Lambert, has, in fact, adopted a similar procedure in a follow-up study of cases transferred to three grammar schools. She obtains multiple correlations for her tests of approximately 0·60 with the academic and technical subjects, respectively. Here, once again, if allowance were made for the fact that the children so transferred had been specially selected, the figure would be raised still higher—how high, it is impossible to tell from the data she gives.

---

[1] " Validating Tests for Personnel Selection," *Brit. J. Psych.*, XXXIV, 1943, pp. 1-19. Worked examples are given from tests used in allocating pupils to trade and secondary (grammar) schools.

Miss Lambert's criterion consisted of eventual "scores in each subject of the grammar school curriculum." This provides a definite criterion that lends itself admirably to a correlational technique. It was the criterion adopted in several of our investigations into the L.C.C. scholarship system and in the investigation organized by the Scottish Council for Research in Education.[1] It provides at once a single criterion and a quantitative criterion. Yet it indicates only the first step in testing the efficiency of our methods of selection. At the age of eleven, as a result of a scholarship examination, I myself was in effect deemed "able to profit" by a secondary education consisting mainly in courses in Latin and Greek. Yet, looking back, I should not feel that the wisdom (or unwisdom) of that allocation could be simply decided by a single weighted total of the marks I obtained in the other examinations some three, five, or ten years later.

One line of approach, which I have ventured to suggest, has not as yet been followed by any investigator. Yet I think it might lead to a useful solution of the main problem. This is the use of what may be called a 'multiple discriminant function.'

Suppose I have applied a single intelligence test (like the Binet scale) to all the pupils in (i) the special m.d. schools and (ii) the ordinary elementary schools of a single borough. I might then be asked: Assuming that Tom comes from a special and Dick from an ordinary school, how would you estimate the most probable value for the intelligence of each? Or conversely, given the actual intelligence of each, to which school is it most likely that he belongs—special or ordinary? In either case what is the accuracy of your answer? Such problems were long ago solved for this simple instance.[2] The method is merely to find a borderline that will maximize the likelihood of our estimates and assignments; and, since we are dealing with a two-fold classification of data into two series, the accuracy will be assessed by what I call a point biserial correlation. The dichotomous classification is given as a mere qualitative distinction; but it can be turned into a quantitative distinction by assigning $+1$ to all the normal children and -1 to all the defective.[3] A little algebra shows that the correlation is then proportional to the standardized difference between the means.

Now suppose that not one, but a number of tests (n) have been applied: I may be asked to calculate from my data the best weighting of those tests, which would enable me to sort Tom, Dick, and all the other pupils whom I may test in future, into (i) the class of mentally defective children and (ii) the class of ordinary or normal children. By the best weighting of the tests I mean that which will yield the smallest number of misclassifications. Once again, the obvious procedure is to maximize the likelihood of my allocations being correct; and this in turn leads, as so often, to a least squares solution.

---

[1] McClelland, W.: *Selection for Secondary Education* (1942). In discussing the choice of his criterion, Dr. McClelland remarks: "A child may profit much more from a secondary course than his success mark would indicate." I would prefer to say "a child may profit, either positively *or negatively*, much more subtly, and much more multifariously, than his success mark may indicate."

[2] The formula reached was given in *Mental and Scholastic Tests*, 1926, p. 165 (1947 ed., p. 177).

[3] The device of quantifying a qualitative classification was first suggested by Yule (*Introduction to Theory of Statistics*, 1912, p. 216), who, however, only employed it for a four-fold table. I first suggested its use for biserial data in dealing with test-results and other measurements for the two sexes (*J. Exp. Ped.*, II, 1912, p. 281). For its relation to other correlation formulæ and the problem of discrimination, cf. Burt, C.: *Psychometrika*, IX, 1944, p. 224, equations 6 and 7, and reference 7.

The multiple regression equation that we are seeking will depend upon (*a*) the n ' point correlations ' with the criterion, and (*b*) the $n \times n$ matrix of intercorrelations. The weighting equation so obtained is identical with what is now commonly known as a ' discriminant function ' ; but it enables us to discriminate two groups or classes only.[1]

However, it can readily be generalized. Instead of two groups we might have three or more, e.g., pupils of special schools, ordinary schools, central schools, and secondary schools, respectively. Such a classification we might regard, under a pre-war scheme of education, as based on variation in a single dimension only (' general intelligence ') ; and on that basis we can easily work out a *multiple* discriminant function. When we add trade schools, we have to take into account a further kind of difference, in a special as well as general ability ; and when we come to the problem of allocation to different types of secondary schools—grammar, technical, and modern, as ordinarily envisaged—we are compelled to assume that the classes cannot be arranged in a single dimension, i.e., along one and the same straight line, but depend upon two or more dimensions or ' factors.' This means a further step in the generalization of the formula. But its derivation in accordance with the usual statistical principles is quite a straightforward piece of matrix algebra. The expression reached is $w'(R_{bb}\,R_{aa}^{-1} - \lambda I) = 0$, where w' denotes the row-vector of weights required, R_{aa} the matrix of total codeviances, R_{bb} the matrix of codeviances ' between groups,' and λ an undetermined multiplier which proves to be the square of the correlation ratio required for testing significance.[2]

This method of determining discriminative weights proved perfectly feasible in tentative investigations on the training-allocations of recruits for the Army ; and it could readily be employed to determine theoretical methods of allocation to different types of secondary schools.

(v) *No Proof of Marked Differences.*—Dr. Bradford gives the " experimenter's point of view." As he states it, his main conclusion appears to be quite unexceptionable. His results, he says, " suggest that measurable differences (between success at verbal and practical tests, respectively) may be expected among pupils of 11 years of age, if their mental age is at least 12 years." Once again all, I think, are agreed that measurable differences " may be expected." But what we want to know is the size of these differences when they have been measured. Here, Dr. Bradford seems more doubtful : " No completely satisfactory proof of *marked* differences has so far been published."

Two other inferences are worth noting. First, he considers that " non-verbal tests cannot be safely substituted for practical tests." Secondly, he thinks that " in the selection procedure the assessment of general or average ability should be made independently of the assessment of bias of ability towards the verbal or the practical extremes." In this connection it should be noted that, in order to obtain independent assessments, it is not essential to use independent tests. Once again military experience is instructive. In the early stages of the work on personnel selection, separate tests of intelligence were included. When sufficient data had been accumulated, however, it became possible to show that, by means of weighting equations (based on partial regressions), the weighted total obtained with a varied set of tests dealing with relatively specialized abilities could also be used to assess general ability. Such a procedure is economical of time, labour, and expense. Nevertheless, until further experience has been secured, I should agree that it would be wiser to include separate tests of intelligence, since they are so well standardized and their reliability and validity are so well known.

---

[1] The earliest attempt to devise a discriminating function is that of KARL PEARSON (" On the Coefficient of Racial Unlikeness," *Biometrika,* XVIII, 1926, pp. 105f.). This procedure (cf. *Psychometrika,* III, 1938, p. 153), modified along the lines suggested above, was used in an investigation on classification of school children for central, secondary, trade and other types of school. Recent interest in the subject dates from Miss Barnard's paper on classifying Egyptian skulls (*Ann. Eug.*, VI, 1935, pp. 352f. Cf. FISHER, R.A. : " The Coefficient of Racial Unlikeness," *J. Roy. Anthrop. Soc.*, LXVI, 1936, pp. 57f. : Cf. *id.*, *Ann. Eug.*, VI, 1937, pp. 179f.). At my suggestion, one of my own research students applied the method to the weighting of psychological tests for a twofold vocational classification : (TRAVERS, R. M. W. : " The Use of a Discriminant Function in Treating Psychological Group Differences," *Psychometrika,* IV, 1939, pp. 25f.).

[2] A simple form of proof is given in roneo'd notes on " The Derivation of the Discriminant Function (1945)," obtainable from the laboratory. An alternative proof from the formula for canonical correlation was briefly indicated in *Brit. J. Psych., Stat. Sect.*, I, 1948, p. 105. It will be noted that the equation given in the text requires us to determine the latent roots of an *asymmetrical matrix* ; the working method we adopt is virtually that described by FRAZER, DUNCAN and COLLAR (*Elementary Matrices,* 1938, p. 142f.) slightly simplified. The number of significant dimensions can be determined by a preliminary factor analysis by least squares. I may add that, so far as computations are concerned, Dr. Hamilton has recently found the method quite practicable for studying diagnostic discriminations between different groups of patients.

Mr. Dempster would apparently agree with Dr. Bradford. Studying the problem from " the selector's point of view," he concludes that " in few cases do either spatial or performance tests indicate clear bias, especially among children with higher intelligence," and that consequently " their practical value in educational guidance is very limited." This, I think, is very true of tests available at present. With further research perhaps more efficient methods may be devised. But it now seems clear that, on the weight of evidence as summarized in all these contributions, it is highly doubtful whether such tests will ever be able to demonstrate any marked bias, latent or developed, in any large number of cases at so early an age as eleven plus.

(vi) *The Study of Practicable Procedures.*—Mr. Rodger, in his short but suggestive paper, concludes the series by stating the industrialist's point of view. As an industrial psychologist, he is naturally interested in what the schools are doing, or proposing to do, with those who will become the industrial workers of the future. His practical experience leads him to note another difference between those who opened the debate. " Alexander," he says, " the adviser on administration, emphasizes the need for objectivity," that is, for a procedure which will be " administratively convenient and politically defensible " : " Burt, the research scientist, stresses the importance of a comprehensive and balanced study of the child." And he evidently feels that there will have to be some measure of mutual concession from both sides.

I fully agree that, as I said in my opening paper, " an ideal plan is almost of necessity an impracticable plan." In all problems of applied science, there are two opposing points of view ; and the policy finally adopted for practical purposes has usually to be a compromise. However, at the outset of a joint discussion, it seemed to me desirable that each specialist should start off by making out a full case for his own line of approach. Then, if anything has to be sacrificed, it can be sacrificed consciously, with a knowledge, or at least a hope, that the minimum essentials are being preserved, when the final balance is struck.

In the present instance, the two standpoints are not so much contradictory as complementary ; and their respective requirements will, I fancy, not prove so irreconcilable as some critics have feared. When I was first appointed educational psychologist to the London County Council, a number of my colleagues warned me that I should find myself impaled on, or at any rate perched between, the sharp horns of a dilemma. As a scientist I should find myself wanting to elaborate tentative techniques which would be too cumbersome and expensive for wholesale use. As an employee of the education authority I should probably be forced to sacrifice scientific accuracy and caution to the need for snapshot methods, which would not only have to be workable and cheap, but would also have to appear intelligible and plausible to the non-scientific members of the Council's committees. My solution was to carry on my researches as an unpaid investigator during the leisure half of my time, and in the other half to apply the knowledge so gained to the practical problems referred to me. In deciding which pupils were, and which were not, mentally defective, the most practicable procedure was to deal with the numerous obvious cases in a quick and summary fashion, and so leave plenty of time for the borderline and problematic cases. That, I think, will prove a useful policy in the allocation of children towards the other end of the scale.

Research is thus needed on two types of problem. So far as theoretical principles are concerned, research in the early stages is bound to be a gamble in the dark. In the early stages, therefore, I think it the wisest plan for investigations to be unpaid. Once the main theoretical problems have been largely solved, then research on practical applications becomes the paramount need. The Hertzes and the Lodges must be followed by the Marconis. Such research should be paid for, either directly or indirectly, by the body which will benefit by the results. Mr. Rodger himself has described the " seven point plan " evolved by the National Institute of Industrial Psychology to assist in problems of vocational guidance and selection; and I readily agree that, with minor modifications and a little further study, it would make an admirable basis for problems of educational guidance and selection. The procedure described in Mr. Dempster's appendix illustrates the working steps by which decisions can be reached first in the more numerous and obvious cases, and then in the rarer and more difficult. Mr. Rodger, however, in his postscript warns us that, " detecting special bents " is in most cases extremely difficult *even at the age of fifteen,* and to attempt to do so at eleven is bound to give " the whole business an air of unreality." Once again, therefore, the primary need is for more knowledge.

(vii) *The Quota Principle.*—Seeing that we are all agreed that standardized tests should form only one part of the basis on which allocations are to be made, and that other information, such as that contained in teachers' reports, ought also to be taken into account, particularly in doubtful cases, we are still left with the difficulty of equating the reports from different schools. Here the ' quota principle ' proposed by Professor Valentine seems to me to provide an admirable solution. According to this plan, the results of standardized intelligence tests determine the *number* of places to be allotted to each primary school; and the actual *individuals* in each school are then selected in accordance with an order of merit, in which the teacher's report may have the same weight as the tests or even greater weight. Mr. V. J. Moore has described how the scheme has been used by one particular authority, and " has received practically unanimous support from the heads of all schools concerned, secondary as well as primary." More recently it has been recommended in the Report of the National Union of Teachers.[1]

May I conclude with the proviso which I tried to emphasize at the outset of the discussion ? To my mind the essential problem is " administrative rather than psychological : psychological considerations must necessarily limit what we decide to be possible, nevertheless psychological conditions alone cannot determine either our aims or our methods."

The psychologist as such is not concerned with ends or aims. Diverse and basic though they may be, the individual's inborn aptitudes and other psychological characteristics can never be the *sole* grounds by which his education is to be decided, though they may frequently limit what is practicable or worth while. Suppose (to take an imaginary case) that the psychologist could demonstrate by 11+ that 30 per cent. of our pupils were distinguished by high musical aptitude, that being the strongest of their natural gifts ; we should not argue that these 30 per cent. should be given a predominantly musical education. In peace-time as in war a civilized community in the twentieth century needs certain fairly definite proportions of its population to do this or that fairly definite type of work. Fortunately, the frequency-distribution of social needs appears to correspond pretty well with the frequency-distribution of psychological capacities and characteristics; further, the educationist may reasonably consider the leisure activities, as well as the industrial activities, of the individual.

---

[1] VALENTINE, C. W. : *Examinations and the Examinee* (1938), p. 11f. National Union of Teachers : *Transfer from Primary to Secondary Schools* (1949).

After all, even before the age of 11+ we find occasional instances of special abilities and disabilities. But, unless these are exceptionally well marked or the child's general intelligence severely limited, we do not argue that in such cases the content of the curriculum should be altered; rather we urge that the methods of teaching should be adapted to the aptitudes, interests, and mental limitations of the child. Similarly, with children who have reached the age of 11+, the psychologist is not justified in forthwith demanding a special curriculum or type of school, unless that is also deemed desirable on other grounds. There are many strong arguments against premature specialization; and in planning any democratic scheme of education, what we have to consider are not merely the personal inclinations of each separate individual, but also the requirements of society as a whole, that is, the requirements of *all* individuals jointly. In this sense, therefore, I would support the policy which Mr. Rodger has urged under the somewhat provocative title of " planned procrastination."

III.—SUMMARY AND CONCLUSIONS.

1.—An ideal plan of selection would take into account, not only innate intelligence, special aptitudes, and educational attainments, but also health, temperament and character, leading interests, and (with due precaution) social conditions.

2.—Of these, general intelligence and educational attainments can be assessed with adequate accuracy by standardized tests.

3.—It is generally admitted that, at the age of eleven plus (mental rather than chronological), some slight tendency towards either an academic or a technical bias is discernible, at least in a small proportion of cases. But most of the contributors plainly doubt whether the number of recognizable cases, showing a strong and permanent bias, will be large enough at this early age to use such a difference as a basis for a broad or final classification on any comprehensive scale.

4.—The requirements of a plausible working scheme suggest that, whatever the age at which such classifications are attempted, the clearer cases may be dealt with summarily and briefly; the borderline cases studied in greater detail; and the cases that are still left in doubt be deferred for ultimate allocation to a later stage.

5.—Psychological considerations may often be limiting factors, but they cannot be the sole factors, in considering the most appropriate form of education for each individual child.

THE LEISURE ACTIVITIES OF GRAMMAR SCHOOL CHILDREN.

By MARY STEWART

(*Tutorial Class Tutor, University of London*).[1]

I.—*Introduction.* II.—*Size of family.* III.—*Periodicals.* IV.—*Newspapers.* V.—*Cinema.* VI.—*Radio.* VII.—*Clubs and associations.* VIII.—*Favourite out of school activities.* IX.—*Comments and conclusions.* X.—*Summary.* XI.—*Questionnaire.*

I.—INTRODUCTION.

This investigation, which is a continuation of that made by us in the Summer of 1946,[2] was undertaken primarily in order to compare the leisure activities of grammar school children with those of children in modern secondary schools. A questionnaire,[3] similar to that which was answered by all modern school children in Ilford, in 1946, was answered by boys and girls in five grammar schools (two boys', two girls' and one mixed) in the Ilford and Wanstead areas. We are again greatly indebted to the Director of Education and to the headmasters and headmistresses who permitted the children to answer the questions in school hours during the Spring term, 1947. Replies were obtained from 675 boys and 801 girls. This number included nearly all the boys and girls between 11 and 15 years and about 300 of those of 15 years and over. Where statistical comparisons have been made between grammar and modern school children, grammar school children of 15 years and over have been excluded.

All percentage differences which are discussed in the text have been checked and have been found to be significant with a critical ratio of not less than 2·0.

II.—SIZE OF FAMILY.

The children were asked to state how many brothers and sisters they had since it seemed reasonable to suppose that their activities were influenced by the size of their families. Child members of small families tend to receive more parental supervision, to have more space at home for private possessions and hobbies, and to have more pocket money than members of large families. We think, therefore, that Table 1, which shews the number of siblings possessed by grammar and modern school children is of considerable interest. On an average, an Ilford grammar school child has 1·2 and a modern school child 2·2 siblings. 28 per cent. of the grammar school children and 17 per cent. of those at modern schools are only children whilst 42 per cent. of the former and 30 per cent. of the latter have either a brother or a sister. Children with four or more siblings are found five times as often in modern as in grammar schools.

---

[1] This report is based upon an enquiry made by the members of an Ilford Tutorial class. The general direction of the enquiry was in the hands of the tutor, Mary Stewart, but the findings and conclusions given in this report are the outcome of much work and discussion on the part of all members of the group. The members of the group were ; Mrs. M. Stewart, Mr. J. Amon, Miss E. Binks, Mrs. H. Dutton, Mrs. E. Fulton, Miss E. Morgan, Mrs. N. Newport, Mrs. M. Owen, Mr. P. Shaw, Miss P. Taylor, Miss D. Telford, Mr. C. Williams, Miss B. Windle.

[2] *The Leisure Activities of School Children* (W.E.A., 1948).

[3] See end of this paper.

TABLE I.

Boys and Girls		
11-14 years, inclusive.		
	Grammar Schools	Modern Schools
Total Children	1476	3046
No. of Siblings 0	% 27·8	% 16·5
1	42·1	30·4
2	18·7	21·2
3	7·4	11·6
4 or more	3·9	20·3

III.—PERIODICALS.

Questions : *Name the comics and adventures (e.g., " Beano," " Girls' Crystal," " Champion," etc., you read regularly for pleasure.*

Name any other magazines and papers (excluding newspapers) you read regularly.

As in our previous investigation,[1] we classified as ' bloods ' or ' adventures ' such papers as *Champion* and *Hotspur* and the *Girls' Crystal*, which consist primarily of school and adventure stories of a fantastic nature. We classified as comics such papers as *Dandy* and *Beano*, which consist primarily of comic strips. Comics sometimes contain stories whilst bloods may give a small amount of space to comic strips. We put the *Girls' Crystal* into the ' blood ' category as it has the same type of appeal for girls as the *Champion* has for boys. Thus a copy of the *Girls' Crystal*, which we read, contained stories about ' The Gipsy Violinist's Warning . . . Why was the Gipsy girl so desperately anxious that Beryl and her chums should not visit the Witch's grotto? . . . ; about ' The Boy who Bosses . . . Forced to take orders from the boy they despised . . . ' ; about ' The Case of the Frightened Girl . . . Accused of being in League with the Green Rajah . . . '

Our investigation, like that of Jenkinson[2] shows the great popularity of bloods and comics among boys and girls, but our figures do not bear direct comparison with his owing to the different methods of classification. Jenkinson took ' bloods ' to mean ' the weekly, fortnightly or monthly adventure and story papers and magazines for boys and girls[2] and included in this category periodicals which we placed in the ' comic ' or in some other class. We found that Jenkinson's method of classification did not show clearly certain differences relating to age and sex and type of school. Thus, unlike Jenkinson who found that " at all ages, about one-quarter of the girls' ' bloods ' reading is devoted to boys' bloods '," we found that girls rarely read boys' bloods such as *Champion* and *Hotspur*. In fact, only 5 per cent. of the modern school girls and 8 per cent. of the grammar school girls read periodicals which we regarded as boys' bloods, whilst 30 per cent. of the modern school and 36 per cent. of the grammar school girls read the *Girls' Crystal*. We found that comics, on the other hand, not only were almost as popular with girls as with boys, but individual comics had the same degree of appeal. Thus *Dandy*, *Beano* and *Film Fun* headed both the girls' and boys' lists. Similarly, comics lost their appeal with advancing years at much the same rate among both boys and girls.

---

[1] *The Leisure Activities of School Children* (W.E.A., 1948).

[2] A. J. JENKINSON : *What Do Boys and Girls Read?* (Methuen, 1940, Ch. 4).

TABLE II.

AVERAGE NUMBER OF 'BLOODS' AND COMICS PER CHILD.

	11 Years.		12 Years.		13 Years.		14 Years.		15 + Years.		11-14 Years inclusive.	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
No. of Children :												
Grammar School	104	68	207	194	166	169	117	139	81	231	675	801
Modern School ..	389	334	465	515	461	437	249	186	—	—	1564	1472
" Bloods " (e.g., *Wizard*) :												
Grammar School	1·9	0·7	1·8	0·7	2·0	0·7	1·4	0·6	0·6	0·1	1·8	0·6
Modern School ..	0·8	0·3	0·9	0·3	1·3	0·4	1·1	0·3	—	—	1·0	0·3
Comics (e.g., *Beano*) :												
Grammar School	1·1	1·0	0·9	1·0	0·7	0·6	0·4	0·3	0·1	—	0·8	0·7
Modern School ..	1·8	1·7	1·6	1·5	1·2	1·1	0·6	0·5	—	—	1·3	1·3
Do not read either " Bloods " or Comics :	%	%	%	%	%	%	%	%	%	%	%	%
Grammar School	12	26	15	19	20	31	42	55	84	92	21	32
Modern School ..	10	19	10	14	12	20	35	49	—	—	15	21

Bloods or adventures were mentioned more frequently than any other form of periodical by grammar school boys in every age group. 46 per cent. of the periodicals named by boys of 11-14 years inclusive, and 21 per cent. of those named by boys of 15+ years were of this type. Each boy read on an average 1·7 bloods, the greatest number being read by 13-year-olds whose average was two per boy.

Next in popularity among the grammar school boys were comics which accounted for 18 per cent. of all periodicals mentioned. These declined in popularity from 27 per cent. of all periodicals named at 11 years to 2 per cent. at 15 years and over. Each boy read on an average 0·7 comics, the greatest number being read by 11-year-olds whose average was 1·1. The percentage of boys who read neither comics nor bloods increased from 12 per cent. at 11 years to 84 per cent. among those of 15 years and over.

Among the grammar school girls, comics were mentioned more frequently than bloods by girls of 11 and 12 years and less frequently by girls of 13 years and over. Among 11-year-old girls, comics and bloods accounted for 35 per cent and 24 per cent., respectively, of all periodicals mentioned. Among those of 15 years and over, these figures had dwindled to 1 per cent. and 5 per cent. respectively.

In every age group, the grammar school boys read more than twice a many bloods as the girls.

In every age group, grammar school boys and girls read more bloods an fewer comics than the modern school boys and girls. The 11-year-old gramma school boys reads more than twice and the 12-year-old twice as many bloods a his opposite number in the modern school. After 12 years, the difference betwee the grammar and modern school children was less pronounced. Bloods reache the peak of their popularity among 13-year-olds at both types of school. N doubt the greater popularity of bloods among grammar school children is due in part, to the fact that whilst all children at a grammar school can read fluentl some modern school children may find it an effort to read a serial or short stor in such a paper as *Champion* or *Hotspur*. Comics make fewer demands o vocabulary and may be appreciated by those who are not fluent readers.

TABLE III.

Periodicals (Excluding Newspapers) Expressed as Percentage Read.

	15+ Years.		11–14 Years inclusive		11–14 Years inclusive.	
	Grammar		Grammar	Modern	Grammar	Modern
	Boys	Girls	Boys	Boys	Girls	Girls
No. of Children	81	231	594	1564	570	1472
No. of Periodicals	214	426	2357	4600	1629	3526
Periodicals per Child	2·6	1·8	3·9	2·9	2·9	2·4
Comics	% 2	% 1	% 20	% 46	% 25	% 55
Bloods	21	5	46	35	23	14
Women's Magazines	1	22	—	—	13	7
Film Magazines	2	9	—	—	3	4
Illustrated Papers	8	25	4	5	9	5
Girls' or Boys' Own	1	4	2	2	4	1
Handicraft	5	—	4	2	—	—
Sport	14	—	6	1	—	—
Youth and Service	5	4	3	2	2	1
Popular Weeklies	8	13	3	1	5	1
Popular Monthlies	3	—	1	—	1	—
Radio Times	5	5	2	—	2	—
Children's Newspaper	1	1	2	—	4	1
Aircraft	5	—	1	2	—	—
Rest	19	11	6	4	9	11

Next to bloods and comics, the most popular periodicals among the grammar school boys were sporting papers which accounted for 6 per cent. of all periodicals read. These periodicals included such papers as *Sporting Record*, *Speedway Gazette*, *Boxing News*, *Football Echo*, and many others. They increased in popularity with age and accounted for 14 per cent. of all periodicals named by boys of 15 years and over. Next in popularity among the boys as a whole came illustrated papers such as *Picture Post* and *Illustrated*.

Among the grammar school girls, women's magazines, such as *Woman's Own* and *Wife and Home* were next in popularity to bloods and comics. Illustrated papers came third and popular weeklies (e.g., *Everybody's*, *John Bull*) came fourth.

IV.—NEWSPAPERS.

Question : [1]*What newspapers, if any, do you read regularly ?*

96 per cent. of the grammar school children as a whole read one or more newspapers regularly. Every girl and all but 1·2 per cent. of the boys of 15 years and over named at least one newspaper. Even among the 11-year-olds, only 3·8 per cent. of the boys and 7·4 per cent. of the girls did not read a newspaper. In many cases, the children may do little more than look at the pictures or scan the headlines, but there is little doubt that among grammar school boys and girls, the habit of at least glancing at a newspaper is well established by, if not before, adolescence.

Each grammar school child read on an average 1·1 national daily papers, 0·5 London evening papers, 0·8 Sunday papers and 0·1 local (mostly weekly) papers. Among the national dailies, the *Daily Mirror*, read by 34 per cent. of the boys and girls was the most widely read with the *Daily Express* (26 per cent.) second, and the *News Chronicle* (11 per cent.), third. The *Evening News* (27 per cent.) was the most widely read evening paper and the *News of the World* (19 per cent.) the most widely read Sunday paper. The last named paper was read by 40 per cent. of the boys of 15 years and over. *The Times* was mentioned by 1 per cent. and the *Daily Telegraph* by 6 per cent. of the children. Only 5 of the 1,476 children mentioned the *Daily Worker*.

Since few school children buy newspapers regularly, it is not surprising to find that the order of popularity of these papers among children is similar to that found among adults in the area in which they live. According to the Hulton Readership Survey, 1949,[2] the most widely read papers in the South-East area of England are the *Daily Mirror* and the *Daily Express*, which are read by 36 per cent. and 25 per cent., respectively, of the adult population. The most widely read Sunday and evening papers in the same area are the *News of the World* and the *Evening News*, which are read by 51 per cent. and 36 per cent., respectively, of the adult population.

The average number of daily (i.e., morning and evening) papers read by the grammar school boys and girls of 13 years and over was 1·7 and 1·6, respectively. The corresponding figures obtained by W. D. Wall,[3] in his investigation among boys and girls (13-16 years) in grammar schools in the provinces in 1944 and 1945 were 1·9 and 1·8. As was to be expected, W. D. Wall found that the papers read by the boys and girls in the provinces included a far higher proportion of local daily papers than we found among the children in our area which is adjacent to London.

---

[1] This question was not included in the questionnaire issued to the Modern School children. See *The Leisure Activities of School Children* (W.E.A., 1948).

[2] Published privately by the Hulton Press, 1949.

[3] W. D. WALL : "The Reading of Newspapers, among Adolescents and Adults, Vol. 18 of this *Journal*, Feb. and June, 1948).

TABLE IV.

NEWSPAPERS

AVERAGE NUMBER OF NEWSPAPERS READ PER GRAMMAR SCHOOL CHILD.
PERCENTAGE OF GRAMMAR SCHOOL CHILDREN WHO DO NOT READ A NEWSPAPER.

	11 Years		12 Years		13 Years		14 Years		15+ Years		All Ages	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Total Children	104	68	207	194	166	169	117	139	81	231	675	801
National Daily Papers..........	1·1	0·9	1·1	1·0	1·2	1·0	1·1	1·15	1·2	1·1	1·1	1·0
London Evening Papers..........	0·5	0·3	0·5	0·5	0·5	0·6	0·5	0·4	0·6	0·5	0·5	0·5
Sunday Papers ...	0·6	0·5	0·7	0·7	1·0	0·9	0·9	1·0	1·1	0·8	0·9	0·8
Local Papers	0·1	0·03	0·1	0·2	0·1	0·2	0·1	0·2	0·2	0·2	0·1	0·2
All Papers	2·3	1·8	2·3	2·4	2·9	2·5	2·6	2·8	3·4	2·6	2·6	2·5
Do not read a Paper	% 3·8	% 7·4	% 4·3	% 6·2	% 4·2	% 7·7	% 5·1	% 2·2	% 1·2	% —	% 4·0	% 4·1

V.—CINEMA.

Question: *How many times a week do you go to the cinema?*
If you do not go weekly, how often do you go?

It will be seen from Table 5 that 30 per cent. of both the boys and girls at grammar schools went to the cinema once a week whilst 8 per cent. of the boys and 3 per cent. of the girls went twice or more each week. The average weekly attendance of these boys and girls is lower than those of the grammar school boys and girls (13 – 16+ years) in the Midlands studied by a research group at Birmingham University in 1946-47.[1] The average weekly attendances of the boys and girls in the latter group were 1·0 and 0·9, respectively, whilst those of our group were 0·59 and 0·49.

It will be seen from Table VI that the average weekly attendance of the modern school boys and girls was almost twice as great as that of the grammar school children of the same age. W. D. Wall[1] and J. Ward[2] and A. J. Jenkinson[3] report also a higher attendance among non-grammar school children. The relatively low attendance at cinemas of grammar school children appears to be due to a number of causes. One is, no doubt, the homework which occupies these children on week nights. Another is a more developed critical sense which caused many children to reply 'When there is a good film on,' in answer to the question, 'How many times a week do you go to the cinema?' This reply was rarely given by modern school children. It is probable, too, that grammar school teachers have, through parent-teacher associations and other contacts, more opportunities than modern school teachers, to urge parents to limit the cinema visits of their children.

J. Ward[2] gives the frequency of cinema visits for children of 5-15 years both in rural and in urban areas, but as the age group is not broken down into separate years or into two or three smaller age groups for rural and urban areas separately, a comparison of our data with that obtained in the Social Survey has only limited value. J. Ward found that among urban children of 5-15 years in modern, technical and grammar schools, 35 per cent. went once, and 11 per cent.

---

[1] W. D. WALL: "The Adolescent and the Cinema" (*Educ. Rev.*, Vol. 1, Nos. 1 and 2 1948).

[2] J. WARD: *Children Out of School* (Social Survey, 1948).

[3] A. J. JENKINSON: *What do Boys and Girls Read?* (Methuen, 1940).

TABLE V.

CINEMA VISITS PER WEEK—GRAMMAR SCHOOLS.

	11 Years		12 Years		13 Years		14 Years		15+ Years		All Ages	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
No. of Children	104	68	207	194	166	169	117	139	81	231	675	801
	%	%	%	%	%	%	%	%	%	%	%	%
3 or more visits	—	—	1	—	1	1	1	1	1	—	1	—
2 Visits	8	6	6	3	7	5	8	2	10	1	7	3
1 Visit	34	29	32	24	30	34	23	35	31	28	30	30
Once a Fortnight ..	6	6	9	14	7	11	13	9	7	15	9	12
Once a Month	21	9	18	15	23	20	13	18	6	16	18	17
Sometimes	29	46	30	41	30	29	39	33	41	37	32	36
Never	2	4	3	3	2	1	3	2	4	2	3	2
Average per Week..	0·6	0·5	0·59	0·44	0·58	0·57	0·54	0·53	0·63	0·45	0·59	·499

TABLE VI.

CINEMA VISITS PER WEEK—GRAMMAR AND MODERN SCHOOLS—11-14 YEARS, INCLUSIVE.

	Only Children.				All Children.			
	Boys		Girls		Boys		Girls	
	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.
No. of Children	180	257	143	244	594	1564	570	1472
	%	%	%	%	%	%	%	%
3 or more Visits.........	1	4	1	1	1	5	—	3
2 Visits	6	19	4	18	7	22	4	18
1 Visit	33	58	34	62	30	53	30	54
Once a Fortnight	7	2	9	1	9	1	11	3
Once a Month	18	1	17	2	19	2	17	1
Sometimes	31	6	31	10	31	6	36	12
Never	4	10	4	6	3	11	2	9
Average per Week	0·57	1·10	0·56	1·03	0·58	1·12	0·51	1·01

went more than once a week. We found that among the grammar and modern school children of 11-15 years, 47 per cent. went once and 19 per cent. went more than once a week. J. Ward found, as was to be expected, that urban children visited the cinema more frequently than rural children and that when urban and rural figures were combined, 53 per cent. of the boys and 45 per cent. of the girls in the 11-15 age group went once or more each week to the cinema compared with 28 per cent. of the boys and 32 per cent. of the girls of 5-7 years and 49 per cent. of the boys and 45 per cent. of the girls of 8-10 years. Our survey confirms the trend of W. D. Wall's data and that of the Social Survey that boys visit the cinema rather more frequently than girls.

In the light of the above figures, the proportion of grammar school boys and girls who belong to Saturday morning Cinema Clubs (see Table X) is surprisingly large. In the age group 11-14 years, inclusive, 30 per cent. of the grammar school and 38 per cent. of the modern school boys belonged to these clubs. The corresponding figures for the girls were 20 per cent. and 26 per cent. As was to be expected, membership of these clubs declined after 12 years. We were interested to note, however, that the proportion of 14-year-old grammar school girls belonging to these clubs was nearly twice as great as that among modern school girls. Probably 14-year-old modern school girls who know that they will shortly be working in office or factory and have money of their own to spend are more attracted by adult leisure pursuits than their opposite numbers in grammar schools who will spend two or more years in school.

It is clear both from this and similar studies[1] that there has been a large increase in cinema attendance among school children since the L.C.C.[2] made its investigation in 1931. During that year, the L.C.C. found that 30 per cent. of the children in its schools went once whilst only 9 per cent. went twice or more each week. We do not think that the cinema is harmful to children in the sense that it encourages anti-social behaviour, but we think it regrettable that modern school children, who are already handicapped by a school leaving age which is rarely less than one year and is often two or three years earlier than that found in grammar schools, should spend so many more hours than their more fortunate brethren, watching films which frequently display poor taste and judgment and which, for the most part, provide only fantasy satisfaction for spectators.

TYPE OF FILM ENJOYED BY GRAMMAR SCHOOL CHILDREN.

Question[3]: *What type of film do you enjoy most? (e.g., Westerns, Musical, Cartoons, Crime, etc.).*

Among the boys, crime films, which were mentioned by more than 50 per cent. of every age group, were more popular than any other type at every age, save 11 years, when they came second to Westerns. Cartoons came second in every age group save that of 11 when they came third. Westerns were third in popularity with the boys as a whole, although their popularity declined with age from 66 per cent. at 11 years to 16 per cent. at 15 years and over. No other type of film was named by as many as 20 per cent. of the boys as a whole.

Girls between 11 and 14 years, inclusive, enjoyed Musicals more than any other type of film. Over that age films variously described as 'Classics,'

---

[1] J. WARD: *Children Out of School* (Social Survey, 1948).
W. D. WALL: "The Adolescent and the Cinema" (*Educ. Rev.*, 1948).
TYNEMOUTH YOUTH COUNCIL: *Youth Survey Report*, 1947.
K. BOX: *The Cinema and the Public* (Social Survey, 1946).

[2] F. H. SPENCER: *School Children in the Cinema* (L.C.C., 1932).

[3] Modern school children were not asked this question.

'Classical' and 'Literary' came first. Films of this type most frequently named were "Great Expectations," Nicholas Nickleby," "Henry V," and "Pride and Prejudice." Crime films were placed second by girls between 11 and 14 years, inclusive. After that age there was a marked decline in the interest shewn in this type of film. Cartoons declined in popularity among the girls from 40 per cent. at 11 years to 12 per cent. among those of 15 years and over.

TABLE VII.

TYPES OF FILM ENJOYED BY GRAMMAR SCHOOL CHILDREN.

	11 Years		12 Years		13 Years		14 Years		15 + Years		All Ages	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
No. of Children	104	68	207	194	166	169	117	139	81	231	675	801
Crime	% 60	% 44	% 64	% 48	% 65	% 49	% 51	% 41	% 57	% 22	% 61	% 39
Cartoons	53	40	52	32	49	19	38	19	39	12	48	22
Westerns..........	66	29	47	30	45	21	35	10	16	6	44	18
Musical	14	51	13	54	16	60	30	54	20	36	18	50
Comedy...........	15	6	11	8	15	9	11	6	18	6	14	7
Adventure	14	7	4	6	2	5	4	4	2	3	5	5
Historical	7	7	2	8	5	10	4	8	4	19	4	12
Documentary Educational	2	1	4	3	5	1	3	8	5	7	4	4
Drama............	2	—	3	6	3	4	3	13	6	23	3	13
'Love'—Romance .	1	9	2	4	5	8	3	9	6	5	3	6
'Classics' 'Classical'.	4	1	2	5	1	9	5	33	5	45	3	22
Nature—Animal ...	5	10	2	11	4	11	3	3	—	2	3	7
Thrillers—Horrific..	3	—	1	5	2	5	2	7	7	3	2	4
British............	—	—	—	1	—	3	—	11	7	9	1	6
Others	11	3	9	5	4	12	4	9	2	5	6	7

Our figures suggest that crime films are the only type which promoters of film shows can be reasonably confident will be appreciated by both boys and girls. The great popularity of Musicals among girls is due, no doubt, not merely to their enjoyment of light music which is reflected in their replies to the Radio programme question, but to their appreciation of light romance, dancing and the gaiety which are usually found in these films. Boys, on the other hand, clearly enjoy more than girls both slapstick and the type of humour found in cartoons. Thus cartoons are mentioned more than twice as often and Comedy films twice as often by boys as by girls.

VI.—RADIO.

Question: *Name the Radio programmes you listen to regularly with pleasure.*

Each grammar school child named on an average 4·5 and each modern school child 3·3 programmes. Nearly 99 per cent. of the grammar school children and 96·5 per cent. of the modern school children were radio listeners.

There were marked changes in taste among boys and girls at grammar schools with the advance of years. Thus the percentage listening to Play and Serials increased between 11 and 15 + years from 18 per cent. to 73 per cent. among the girls and from 9 per cent. to 46 per cent. among the boys. The percentage listening to Light Orchestra increased during the same period from 13 per cent. to 57 per cent. among the girls and from 2 per cent. to 32 per cent. among the boys. The percentage of boys listening to Dick Barton declined from 49 per cent. at 11 years (when it headed the list) to 31 per cent. among those of 15 years and over.

Since the modern school children answered the questionnaire in 1946 and the grammar school children in 1947, a direct comparison can only be made in a few cases. Of those programmes which continue from year to year, Children's Hour was more popular among the grammar school than the modern school children in every age group save that of the 14-year-old girls among whom 24 per cent. of the modern school and 22 per cent. of the grammar school girls listened to it regularly. This difference is due, no doubt, largely to the fact that Children's Hour makes more intellectual demands upon the listener than many other types of programme. Plays were no doubt more popular among grammar school children for the same reason. On the other hand when Variety programmes were expressed as a percentage of the total programmes mentioned, they were found to be almost as popular among grammar school as among modern school boys, the percentages being 35 and 36, respectively. The corresponding figures for grammar school and modern school girls were 29 per cent. and 35 per cent., respectively. J. Ward[1] found that in urban and rural areas combined, variety was equally popular with boys at both grammar and modern schools. She says however, that 'rather unaccountably' it was less popular among modern school than among grammar school girls.

It will be seen from Table VIII that the Children's Hour is more popular with boys than with girls in both types of school. Our figures resemble fairly closely those obtained by J. Ward[1] in urban and rural areas combined. We found that 21 per cent. of the boys and 34 per cent. of the girls in grammar schools and 14 per cent. of the boys and 27 per cent. of the girls in modern schools listened to the Children's Hour. The corresponding figures given by J. Ward are 17 per cent. and 36 per cent. for boys and girls, respectively, in grammar schools and 16 per cent. and 26 per cent. for boys and girls in modern schools. That over one-third of the grammar school and over one quarter of the modern school girls listen regularly to the Children's Hour, reflects great credit on those responsible for this programme. It is a pity, however, that similar success cannot be achieved with the boys. Our results suggest that the programme is not sufficiently vigorous for Dick Barton lovers and that if more time were devoted to adventures in history and fiction and real life, more boys would enjoy the programme.

We consider that the B.B.C. caters adequately as a rule both for the entertainment and the education of children. We do not agree with those who suggest that Dick Barton is partially responsible for the increase in juvenile delinquency. Apart from the fact that the increase in juvenile delinquency started before the

---

[1] J. WARD: *Children Out of School* (Social Survey, 1948).

TABLE VIII

RADIO.

	11 & 12 years		13 & 14 years		15+ years		11–14 years inclusive		11–14 years inclusive	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Boys	Girls	Girls
	Grammar		Grammar		Grammar		Gram.	Mod.	Gram.	Mod.
Total Children	311	262	283	308	81	231	594	1564	570	1472
Ignorance is Bliss ..	% 51	% 32	% 58	% 42	% 57	% 40	% 54	% —	% 37	% —
Twenty Questions ..	43	58	21	53	33	42	42	—	55	—
Dick Barton	40	30	27	17	31	13	33	—	19	—
Paul Temple	39	30	30	20	27	8	35	18	24	22
Itma	29	18	27	24	22	15	28	49	21	38
Merry Go Round ..	28	16	30	33	26	30	29	12	25	12
Children's Hour ...	26	43	16	26	7	10	21	14	34	27
Much Binding	26	17	25	22	20	23	26	—	20	—
Sport	25	1	19	9	24	8	22	4	6	—
Have a Go	24	39	16	23	17	24	20	—	31	—
Radio Forfeits	23	30	25	32	7	12	24	—	31	—
Plays and Serial Stories	20	32	32	58	46	73	25	44	46	46
Just William	13	10	2	6	—	2	8	70	8	72
Daring Dexters	8	12	4	7	—	1	2	—	9	—
Monday Night at 8	7	8	5	4	3	5	6	—	6	—
Carroll Levis	7	14	5	9	1	7	6	19	11	41
Appointment with Fear............	6	2	9	4	14	2	8	33	3	18
Taxi.............	6	3	4	2	1	—	5	—	2	—
Mystery Play House	5	3	8	9	12	3	7	—	6	—
Light Orchestra....	5	13	15	34	32	57	10	6	24	14
News, Talks	5	2	6	2	9	9	5	1	2	1
Quiz..............	4	10	2	11	6	11	3	4	11	5
Serious Music	3	3	4	6	7	15	3	—	5	—
Other Variety	14	13	9	15	1	3	7	—	7	—
Rest..............	22	15	21	21	22	26	21	—	18	—
Do not listen to radio	0·6	1·1	2·1	2·0	1·2	—	1·3	4·3	1·6	2·7

Dick Barton era and that large numbers of well-behaved children greatly enjoy this programme, it is also true that many other forms of entertainment provided for and by children of all ages depict anti-social behaviour. If Dick Barton is unsuitable for children, then, to be logical, we should censor, too, Walt Disney cartoons, Grimm's fairy tales, Shakespeare's plays, Tales from the Odyssey, comics, bloods and many children's games improvised by themselves. We agree, however, that 6.45 p.m. is not usually a convenient hour for children with homework to listen to a radio programme and that many parents would be happier if this, or similar programmes, with a special appeal to children, were reserved for Fridays and Saturdays.

VII.—CLUBS AND ASSOCIATIONS.

Questions answered by Grammar school children.

(a) *Underline any of the following to which you belong: Scouts, Guides, Cinema Club, Boys' or Girls' Brigade, Church, Chapel, Sunday School, Pre-Service Unit, Youth Club, Social Club.*

(b) *Name any other clubs or circles not mentioned above to which you belong.*

Question answered by modern school children.

Name any movement, club or circle (e.g., Girl Guides, Scouts, Church, Chapel, Cinema) to which you belong.

Religious Organisations.—The replies to our questionnaires suggest that membership of a religious organisation is far more common among grammar school than among modern school children. James and Moore[1] in their investigation among adolescents in a working-class district in Manchester in 1939, found also that secondary school boys and girls gave more time to religious activities than children of the same age in elementary schools. The differences we found, however, were so striking that we wondered if the different form in which the questions were worded in the two questionnaires influenced the children's replies. It will be seen above that in each questionnaire, Church and Chapel were mentioned explicitly.

Table IX shows that at 11 years, 73 per cent. of the boys and 68 per cent. of the girls at grammar schools belonged to some religious organisation compared with only 22 per cent. of the boys and 47 per cent. of the girls at modern schools. At 14 years, the figures were 33 per cent. and 68 per cent., respectively, for boys and girls at grammar schools and 12 per cent. for both boys and girls at modern schools. It will be noted that whilst among boys at each type of school and among girls at modern schools, membership declines with age, among grammar school girls membership fluctuates between 59 per cent., and 68 per cent. 24 per cent. of the boys and 65 per cent. of the girls of 15 years and over belong to some religious organisation.

L. J. Barnes states that " the trend of recent systematic investigation is undoubtedly to support the belief that forms of worship have commanding attraction and significance for only a small minority of young people to-day."[2] Judging by our membership figures, this seems to be true of boys and girls of 14 years and over in modern schools and untrue of adolescents at grammar schools. But attendance figures at Church or Chapel or Sunday School may not

---

[1] JAMES, H. E. O., and MOORE, F. T.: " Adolescent Leisure in a Working-Class District " (*Occup. Psych.*, Part I, Vol. 14, 1940, and Part 2, Vol. 18, 1944).

[2] BARNES, L. J.: *Youth Service in an English County* (King George's Jubilee Trust, 1945, p. 93).

TABLE IX.

MEMBERSHIP OF RELIGIOUS ORGANISATIONS—GRAMMAR AND MODERN SCHOOL CHILDREN.

	11 Years				12 Years				13 Years				14 Years			
	Boys		Girls		Boys		Girls		Boys		Girls		Boys		Girls	
	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.	Gram.	Mod.
No. of Children	104	389	68	334	207	465	194	515	166	461	169	437	117	249	139	186
Church or Chapel[1]	% 37	% —	% 23	% —	% 36	% —	% 31	% —	% 28	% —	% 38	% —	% 33	% —	% 55	% —
Sunday School[1]	60	—	62	—	44	—	55	—	29	—	43	—	18	—	35	—
Other Religious Organisations	3	—	6	—	—	—	5	—	—	—	5	—	4	—	4	—
Belong to at least one Religious Organisation	73	22	68	47	57	20	65	33	51	18	59	25	33	12	68	12

---

[1] In classifying the replies of the modern school children, Church, Chapel and Sunday School were put under the same heading as we found that the children did not always distinguish between them,

be a true reflection of the interest that young people take in religion. Thus W. D. Wall[1] reports that among 196 boys and girls (age 14-17 years) of average intelligence, 34 per cent. of the boys and 38 per cent. of the girls said that they had experienced an increased interest in religion since the age of 13. Among 220 university graduates, C. W. Valentine found as many as 78 per cent. of the men and 74 per cent. of the women reported an intensification of religious interest during adolescence.[2]

The relatively high membership among grammar school children is due, no doubt, primarily to the encouragement given to them by their parents. It will be recalled that the average sized family, to which these children belong, is decidedly smaller than that of children attending modern schools. Further, it is still true that the average income level of the parents of grammar school children is higher than that of those whose children attend modern schools, whilst the average home in the former group is better equipped both from a material and a cultural point of view. We think that for these reasons, it is usually more difficult for parents of modern school children to take their offspring to Church or Chapel on a Sunday morning or evening whilst the older children in such families may be required to help in the house or take care of younger brothers or sisters at all times of day. It is possible, too, that ' bookish ' children, of whom a greater number are to be found at grammar schools than at modern schools, may find Sunday School more attractive than those whose interests are primarily of a practical nature. It has also been suggested to us by parents that such children receive more encouragement than the rest from outside sources since they often become the leaders of groups or circles. Without committing ourselves about the merits of Sunday Schools and other religious organisations, we think it unfortunate that the proportion of children belonging to them should differ greatly in different types of schools. There is no evidence to suggest that the benefits to be obtained from attendance at Church or Chapel or Sunday School are dependent on economic status or level of intelligence or scholastic ability. In this sphere, what is desirable for one set of children should be equally desirable for the rest.

Scouts, Guides and Brigades.—35 per cent. of the boys and 32 per cent. of the girls of all ages in grammar schools belonged to one or other of these uniformed organisations. As was to be expected, membership tended to decrease with age, but even among those of 15 years and over, 22 per cent. of both boys and girls remained members.

Differences in the size of the membership of these uniformed organisations among grammar and modern school boys were not great. Between 11 and 14 years, inclusive, 37 per cent. of the former and 34 per cent. of the latter were members. The difference was greatest at 14 years when 35 per cent. of the grammar school and 25 per cent. of the modern school boys were members. Among the girls between 11 and 14 years, 38 per cent. of those at grammar schools and 31 per cent. of those at modern schools belonged to one of these organisations. The difference was again most marked at 14 years, when the membership was more than twice as great (29 per cent. and 12 per cent.) among the grammar school girls.

The higher membership among the older grammar school children is probably due in part to the attitude of their parents. But perhaps an even more important factor is that modern school children, owing to an earlier school

---

[1] WALL, W. D. : *The Adolescent Child* (Methuen, 1948).

[2] C. W. VALENTINE : " Adolescence and Some Problems of Youth Training " (This *Journal*, 1943, Vol. XIII, p. 58.).

TABLE X.

MEMBERSHIP OF CLUBS AND ASSOCIATIONS.

	11 and 12 Years		13 and 14 Years		15+ Years		11–14 Years inclusive		11–14 Years inclusive	
	Grammar		Grammar		Grammar		Gram.	Mod.	Gram.	Mod.
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Boys	Girls	Girls
No. of Children	311	262	283	308	81	231	594	1564	570	1472
	%	%	%	%	%	%	%	%	%	%
Sunday School	52	56	25	40	6	21	39	—	48	—
Church, Chapel[1]	35	29	30	46	21	61	33	—	39	—
Other Religious Organisations[1]	1	5	2	8	—	10	1	—	7	—
Member of one of above Religious Organisations	63	70	44	63	24	65	54	19	64	31
Cinema Clubs	38	23	20	17	6	—	30	38	20	26
Scouts or Guides	33	36	29	29	18	18	31	24	32	25
Brigades	5	10	7	4	4	4	6	10	6	6
Hobbies Clubs	25	2	15	—	9	—	20	1	1	0·3
Youth Clubs	11	11	17	16	25	31	17	12	19	19
Social Clubs	4	8	5	7	14	6				
Athletic Clubs	8	8	7	10	14	12	8	2	9	1

---

[1] See note under Table 9.

leaving age, develop adult tastes in relation to dress and leisure activities sooner than children in grammar schools. To a boy or girl about to leave school and to earn money, these organisations may seem childish. It will be interesting to note whether the raising of the school leaving age modifies the attitude of the older modern school children towards the uniformed organisations.

Hobbies Clubs.—These clubs included stamp, chess, railway, aeromodelling, Meccano and others. Nearly all were organised inside grammar schools and seemed remarkably popular among the younger boys. Between 11 and 15 + years, membership declined from 31 per cent. to 9 per cent. among the boys. Girls rarely belonged to such organisations. Hobbies clubs of this type were hardly mentioned by modern school boys. We think that clubs fostering these interests in modern schools would give much satisfaction since railways, stamps, Meccano, etc., appeal to nearly all boys at certain ages, irrespective of the type of school to which they belong.

Other Clubs and Organisations.—School dramatic, music, scientific and other such societies were mentioned far more frequently by grammar than by modern school children. As with Hobbies Clubs, we think that some (if not all) of these societies would be appreciated as much by modern as by grammar school boys and girls and we regret that the former have fewer opportunities than the latter of belonging to them.

TABLE XI.

FAVOURITE OUT-OF-SCHOOL ACTIVITIES MENTIONED BY 10 PER CENT. OR MORE OF GRAMMAR SCHOOL BOYS OR GIRLS IN ANY AGE GROUP.

	11 Years		12 Years		13 Years		14 Years		15+ Years		11-14 Years inclusive		11-14 Yea inclusiv	
	Grammar		Grammar		Grammar		Grammar		Grammar		Gram.	Mod.	Gram.	Mo
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Boys	Girls	Gi
No. of Children	104	68	207	194	166	169	117	139	81	231	594	1564	570	147
	%	%	%	%	%	%	%	%	%	%	%	%	%	%
Cricket	67	10	69	7	70	4	57	4	23	—	67	49	—	
Football	51	—	65	1	59	1	52	—	32	1	58	78	—	—
Reading	38	73	41	72	34	64	36	68	22	55	38	37	69	58
Swimming	28	68	36	68	34	76	36	69	42	55	34	38	71	36
Watching Sport	25	—	19	3	35	5	35	4	43	6	28	—	—	—
Cycling	23	28	33	39	39	50	38	52	59	55	34	42	44	34
Collecting	22	19	25	20	23	27	16	23	16	5	19	12	23	3
Handwork, Woodwork	13	1	17	4	20	2	22	1	22	2	19	31	2	—
Train Spotting	11	—	8	—	7	—	7	—	5	—	8	—	—	—
Radio	8	19	7	16	14	11	8	11	17	7	9	5	14	5
Cinema	7	13	6	14	7	14	15	19	12	17	8	13	15	24
Tennis	7	7	8	24	12	34	24	45	32	56	12	3	30	5
Walking	6	9	5	14	3	24	8	27	9	34	5	6	20	20
Street Games	5	13	3	6	2	2	3	3	1	—	3	15	5	56
Gardening	5	13	1	6	8	10	7	19	5	6	5	16	11	16
Skating	4	12	1	4	3	6	8	8	4	6	4	9	—	6
Indoor Games	4	3	10	10	12	5	16	4	25	3	11	11	5	3
Art	4	16	5	15	4	16	6	9	—	6	5	15	14	14
Playing Musical Instrument	3	18	3	11	3	10	3	20	4	16	3	—	14	—
Singing	—	—	—	—	—	—	—	—	—	—	—	2	—	13
Listening to Music	3	7	6	12	2	14	8	14	1	16	4	—	12	—
Clubs—Guides, Scouts	3	10	1	9	8	4	11	18	4	9	5	5	10	5
Pets	1	—	1	4	10	6	2	4	1	1	3	4	4	4
Theatre	1	1	—	5	1	6	3	13	4	26	1	—	7	—
Friends	—	1	4	5	11	7	5	4	6	3	4	1	5	4
Dancing	—	7	—	4	1	6	1	14	7	24	—	1	7	14
Needlework	—	18	—	20	—	21	—	19	—	17	—	—	20	31
Knitting	—	17	—	11	—	12	—	16	—	7	—	—	14	—
Organised Games per Child	1·4	0·4	1·5	0·6	1·5	0·6	1·4	0·8	0·9	0·7	1·5	1·4	0·6	0·4

VIII.—FAVOURITE OUT-OF-SCHOOL ACTIVITIES.

Question : *Name the things you like doing best outside school.*

The grammar school boys named on an average 4·5 and the girls 5 activities in reply to the above question. These included ' making telephones,' ' catching snakes,' ' finding ants' nests,' ' thinking and philosophising,' ' travelling with my parents,' ' paper round,' in addition to a large number of more widely shared occupations.

Marked differences were found between boys' and girls' interests. The most popular activities among grammar school boys were cricket (61 per cent.), football (55 per cent.), cycling (37 per cent.), reading (36 per cent.), and swimming (35 per cent.). The five most popular activities among the girls were swimming (66 per cent.) cycling (47 per cent.), tennis (38 per cent.), and walking (24 per cent.). Swimming and reading were mentioned nearly twice as often whilst walking was mentioned approximately four times as often by girls as by boys.

Cycling increased in popularity among the grammar school boys from 23 per cent. to 59 per cent. and among the girls, from 28 per cent. to 55 per cent. between 11 and 15+ years. Watching sport was remarkably popular among the boys. It was mentioned by 43 per cent. of the boys of 15 years and over and even at 11 years was mentioned by one quarter of the boys. Handicrafts were mentioned by between 13 per cent. and 22 per cent. of the boys of each age group.

Collecting, which was mentioned by 18 per cent. of the grammar school boys as a whole, was most popular among the 12-year-olds of whom a quarter were collectors of one or more of a wide variety of objects. 17 per cent. of the girls were collectors, the activity being most popular at 13 years. There was a marked sex difference in the objects collected. The most coveted possessions among the boys were stamps, football programmes, cricket score cards, pictures of sportsmen, knives and coins. Among the girls, the most coveted were pictures of film stars and stamps.

In Table XII we have adopted a method of classification similar to that used by J. W. Reeves and P. Slater in their investigation of the leisure pursuits of members of the A.T.S.[1] It will be seen that 61 per cent. of all the activities named by the boys and 45 per cent. of those named by the girls were in the ' sports ' category. A further analysis showed that two-thirds of the sports activities of the girls, but only one-third of those of the boys were of a non-competitive type. ' Passive ' sporting activities such as watching football matches, speedway, etc., when expressed as a percentage of all activities (sport and non-sport) increased among the boys from 6 per cent. at 11 years to 9 per cent. at 15 years and over. ' Expressive ' activities such as acting, writing, etc., represented 7 per cent. of the girls' and 2 per cent. of the boys' activities whilst non-participant entertainment such as the theatre, cinema, etc., represented 11 per cent. of the girls' and 5 per cent. of the boys' activities.

Football was decidedly more popular among modern than among grammar school boys, whilst the reverse was true of cricket. It may be relevant to note that in a recent Gallup poll, cricket was found to be more popular in the higher than in the lower economic groups. Tennis was mentioned four times as often by grammar school boys and six times as often by grammar school girls, as by boys and girls in modern schools. Collecting was mentioned seven times as often by grammar school as by modern school girls. On the other hand, gardening and

---

[1] J. W. REEVES and P. SLATER; " Age and Intelligence in Relation to Leisure Pursuits (*Occup. Psych.*, Vol. XX, No. 3, 1947).

TABLE XIIA.

OCCUPATIONS EXPRESSED AS PERCENTAGE OF TOTAL ACTIVITIES.

Age	Sex	Number of Interests per Child	Sport				Indoor Games	Reading Study
			Passive e.g. Watching Sport	Competitive e.g. Football	Non-Competitive e.g. Cycling	Total		
			%	%	%	%	%	%
11 Years	Boys	3·95	6·3	38·9	18·8	64·0	1·2	10·5
	Girls	4·53	—	11·4	27·3	38·7	0·7	16·2
12 Years	Boys	4·36	4·3	37·9	20·0	62·2	2·3	9·7
	Girls	4·73	0·6	14·9	27·8	43·3	1·7	15·3
13 Years	Boys	4·67	7·5	33·8	18·2	59·5	2·6	7·2
	Girls	5· 2	1·0	12·6	32·4	46·0	0·9	12·3
14 Years	Boys	4·94	7·3	30·6	21·7	59·6	3·3	8·5
	Girls	5·74	0·8	14·8	28·5	44·1	0·8	11·9
15+Years	Boys	—	9·2	22·1	28·6	59·9	5·2	5·5
	Girls	4·9	1·2	15·1	31·8	48·1	0·5	11·3
All Ages	Boys	4·5	6·6	33·7	20·7	61·0	2·8	8·4
	Girls	5·0	0·8	14·2	30·0	45·0	0·9	12·9

TABLE XIIB.

Age	Sex	Con-structional	Expressive	Dancing	Social	Collecting	Non-Participant Entertainment	Rest
		Handwork	Acting, Writing, etc.		Clubs, Scouts, etc.		Theatre, Cinema, Radio, etc.	
		%	%	%	%	%	%	%
11 Years	Boys	3·4	2·2	—	0·7	8·5	4·6	4·9
	Girls	8·1	8·8	1·6	2·6	4·2	9·1	10·0
12 Years	Boys	5·1	2·6	—	1·1	5·6	4·4	7·0
	Girls	7·5	6·2	0·8	2·8	4·1	9·9	8·4
13 Years	Boys	5·2	1·4	0·1	1·9	6·5	5·0	10·6
	Girls	7·0	6·9	1·3	2·2	5·2	8·6	9·6
14 Years	Boys	5·0	1·9	0·2	3·3	4·7	6·8	6·3
	Girls	6·4	7·4	2·5	3·9	4·0	10·0	9·1
15+Years	Boys	5·2	1·1	1·5	2·1	3·4	7·4	8·7
	Girls	5·5	6·2	4·9	2·4	1·0	13·5	6·6
All Ages	Boys	4·9	1·9	0·3	1·8	5·7	5·4	7·8
	Girls	6·7	6·8	2·4	2·8	3·5	10·6	8·4

cinema were mentioned much more often by boys and girls and dancing by girls in modern schools.

Activities not given in Table XI, as they were not named by 10 per cent. of grammar school boys or girls of any age group, included fishing (modern school boys, 15 per cent. ; grammar school boys, 5 per cent.), playing ball (modern school boys, 13 per cent. ; grammar school boys, negligible), skipping (modern school girls, 10 per cent. ; grammar school girls, 4 per cent.), cooking (modern school girls, 11 per cent. ; grammar school girls, 6 per cent.).

The greater popularity of cooking and housework among modern school girls seems to be due to the more practical interests of these girls and to the fact that their mothers have more need of their services. Collecting on the other hand is a more suitable pastime for grammar school children as they have more room to store their treasures.

IX.—COMMENTS AND CONCLUSIONS.

The main purpose of our research was to obtain factual information. Some readers may not agree with the comments which follow but we think that they are worthy of consideration.

In a survey of secondary school children's leisure activities, it is important to classify the children not only according to age and sex, but according to the type of secondary school to which they go. A statement that the average secondary school child, in an area such as Ilford, goes to the cinema approximately once a week, may mislead if it is not qualified by the statement that the average grammar school child goes not more than once a fortnight whilst the average modern school child goes rather more than once a week. Similarly, a statement that 34 per cent.of boys and girls at secondary schools belong to some religious organisation does not reveal the fact that among grammar school children, the membership between 11 and 14 years, inclusive, is 59 per cent., whilst among modern school children it is 25 per cent. Again it is true that periodicals known variously as comics, bloods and adventures are more popular than any other type among boys at both modern and grammar schools. But a further analysis shows that among modern school boys, comics are vastly more popular than bloods whilst the latter are preferred by grammar school boys. This is more than a matter of academic interest. Comics may be enjoyed by the semi-literate. Bloods demand for their enjoyment, literacy and a relatively large, if unconventional, vocabulary.

Our results suggest that the advantages of grammar school education are not limited to school hours. It is difficult to assess accurately how far these additional advantages are the direct result of the grammar school curriculum and how far to home guidance, to the good intelligence of the children and to other factors. Hobbies' clubs, dramatic, musical and other school societies, which are mentioned often by grammar school and rarely by modern school children, are clearly the result in part of the initiative and hard work of headmasters and headmistresses and their assistants at grammar schools. On the other hand, similar efforts made among children of lower intelligence might have less productive results. A temperate, critical attitude to films, which prevents a grammar school child from paying a weekly visit to the cinema, irrespective of the programme offered, may be due as much to the attitude of his parents as to the training he receives at school. Again, the greater interest taken by grammar school children in Scouts, Guides, the Children's Hour on the radio, Church, Chapel, Sunday School, playing musical instruments, collecting, etc., is probably due to complex factors which include the economic position and attitude of parents, the level of intelligence of the children in addition to school discipline and curricula which encourage concentration and independent action outside, as well as in school.

As we have said, children in modern schools, compared with those in grammar schools, belong not only to larger but to poorer families. Their parents have less time—and perhaps less ability—to give to their home training. They have less room space for hobbies and they have fewer opportunities to concentrate on a radio programme or a story book. They have many more inducements to play on the streets or go to the movies which provide at a moderate price, warmth, comfort and easy entertainment. The result is that modern school children, who need more urgently than those at grammar schools to increase their vocabularies through reading, to develop their powers of criticism through a discriminating selection of radio and cinema programmes, to develop powers of concentration and independence through hobbies, learning to play musical instruments, etc., have fewer opportunities for these activities than their more fortunate brethren.

Obviously the problem of the modern school child's lesiure, which concerns more than 70 per cent. of the child population and which is related closely to housing, education and other national problems, cannot be solved overnight. But it is important to define the problem as accurately as possible before trying to solve it. Compared with those at a grammar school, the activities of a modern school child lack direction. The modern school child has little homework to encourage him to be self-reliant and to sit at a task until it is completed. He is usually without the spur of outside examinations or the knowledge that if he does not reach a recognised scholastic standard, he will not obtain the job that he urgently wants. He lacks usually that superior mental ability which sometimes enables its possessor to enjoy creative work in the scientific or literary or mechanical sphere in his spare time. He all too frequently lacks that easy familiarity with the written word necessary for the enjoyment of literature or, for that matter, the daily paper or an adventure book. An adequate provision of playing fields, recreation rooms, etc., cannot alone solve this problem. Urgently as these are needed, they provide only a foundation upon which to build a structure which will enable the child to lead a happy, useful, stable life. It is not unusual to hear elderly men say that they believe that, as children, they were happier than the modern child in spite of exploitation of children in the labour market and living conditions which were sometimes too terrible to be imagined by the child of to-day. Eighty years ago, the average child in the slum area of a large town attended school irregularly whilst he almost certainly earned coppers to buy himself food or sweets or treats or to help his mother with necessities at home. Judged by modern standards, he was rough, ignorant, uncouth, unhealthy, ill-clad and perhaps anti-social. But unlike the modern child, he needed and used his wits to keep himself alive and apparently enjoyed doing so. There is a danger to-day that in our proper desire to protect children from the evils of this world, and to give them an environment in which they can develop freely along individual lines and are not repressed or crippled by neuroses, that we may deprive them of the opportunity to use their minds to the full extent.

Certain tools are recognised as essential to an adult if he is to be a success both in his work and his private life. One tool is literacy which an exceptionally bright child can start to acquire at three or four, an average child at six and the rest, save for a very small proportion of mentally and physically defective, a year or two later. There is no doubt that the proportion of children in this country who are semi-literate is decidedly higher than it should be. Whilst school is normally the place in which a child learns to read and write, out of school activities can aid or hinder his progress. Parents can help by making it clear to children how much they lose if they cannot read or write with ease and enjoyment ; by giving them appropriate picture and story books from an early age and by having as many books as possible in the home. Public libraries, which like those at Ilford, have attractive sections for children, and librarians who are eager to help and advise young readers or would-be readers, can help greatly too. On the other hand, children's comics, newspapers whose reading matter is mostly in the form of headlines, periodicals whose chief attraction is cartoons, films which can be followed by those who cannot read are no aid to literacy whatever other virtues they may possess.

Another tool of the greatest value to adults is the capacity to sacrifice an immediate pleasure for the sake of a future gain. Grammar school girls may not enjoy sitting indoors doing their homework when they can hear the voices of modern school children playing in the garden or road. Grammar school boys may not enjoy leaving their play and washing their hands and tidying themselves in preparation for Sunday school. But most of these boys and girls obtain a return for their efforts in knowledge gained and habits which are useful to them as adults.

A third tool is discrimination. The radio undoubtedly aids this in that it provides a wide choice of programmes that are suitable for children. Radio plays and serials stories of the *Jane Eyre* type, which are usually very popular with children, who have an opportunity to hear them, help to develop the child's taste and vocabulary and encourage him to read. Children's Hour, which is one of the best programmes produced by the B.B.C. to-day, would probably be as popular among modern school as among grammar school girls, if they were given the same encouragement to listen in at 5 o'clock. This programme should be modified so that it appeals as much to boys as to girls. On the other hand, the cinema rarely aids discrimination. In spite of the concern felt by those interested in the influence of the cinema on children's taste, the proportion of mediocre and sometimes rubbishy films produced still outnumber those which have both entertainment and cultural value, although there are more of the latter than there were two years ago.

Lastly, there is no doubt that adults find that hobbies in which they display proficiency, add greatly to the enjoyment of their leisure hours. Many of these hobbies, which include carpentry, gardening, metal work, stamp collecting, bird-watching, photography, etc., are of a great interest to children at certain stages of their school careers. Handwork and sometimes gardening have their place in the modern school curriculum and, to a lesser extent, in that of grammar schools, but since school days are too short for the child to learn all that he requires for useful citizenship in the modern world, more opportunities and inducements should be given to children to develop their skills in their spare time. We make, therefore, the following suggestions.

(1) That every appropriate local authority should appoint a *Voluntary Committee* whose members would include representatives of parents, teachers and organisations providing for the out-of-school activities of children ; that the sole duty of this Committee should be to plan and supervise these out-of-school activities. That, in the selection of appropriate L.A's under whom these Committees would work, the importance of personal knowledge of and interest in the area should be borne in mind. In the London area, for example, the Borough Council would seem more appropriate than the L.C.C., although the Education Committee of the latter should be represented on the Committee.

(2) That this Committee should attach the highest importance not only to an adequate provision of playing fields and recreation centres, but to literacy, to interests involving active participation and to proficiency in one or more hobbies.

(3) That, through this Committee, organisations such as Scouts, Brigades, Sunday Schools, Children's Sections of Public Libraries, etc., should have more contact with each other, not in order to deprive them of their present freedom or traditional activities, but so that ideas may be shared and a maximum membership attained.

(4) That every head teacher should be provided with a complete list of organisations catering for the out-of-school activities of children and that this list should be placed in every classroom.

(5) That modern secondary school children should not be deprived of the incentive to hard work, ambition and self discipline which is provided for the grammar school child by homework, examinations, and the prospect of obtaining a satisfactory job on leaving school if he works hard during his school days. That if, owing to home conditions and the level of intelligence of the children, homework is considered undesirable, other out-of-school tasks involving craft work, needlework, play acting, puppets, gardening, etc., should be devised and the children's co-operation sought through competitions, prizes and trophies; that these competitions should be held frequently in each area under the direction of the Committee and that all the appropriate children's organisations should be encouraged to participate.

(6) That the differences in cinema attendance of children in different types of secondary school should be brought to the notices of parents in each area, and that a prolonged effort should be made to reduce the number of modern school children (24 per cent. in Ilford and similar areas), who go to the cinema more than once a week.

(7) That public libraries should be given every encouragement (financial and otherwise) to experiment with new ideas in order to make the Children's Section an important cultural centre in the life of a school child. That a well-equipped library should be provided in all modern secondary schools with the least possible delay.

(8) That to forward the above schemes, special grants should be given to local authorities and should be earmarked for the out-of-school activities of children.

X.—SUMMARY.

Size of Family.—A grammar school child had, on an average, 1·23 and a modern school child 2·25 siblings. 27·8 per cent. of grammar school and 16·5 per cent. of modern school children were only children. Children with four or more sibs. were found five times as often in modern as in grammar schools.

Periodicals.—65 per cent. of all the periodicals mentioned by grammar school boys were either bloods or comics. The proportion decreased from 71 per cent. among the 11-year-old boys to 23 per cent. among those of 15 years and over. Among grammar school boys, bloods were much more popular than comics in every age group. Modern school boys, on the other hand, read approximately twice as many comics as bloods at 11 and 12 years. In every age group, grammar school boys read more bloods and fewer comics than their opposite number in modern schools. At 11 years, they read twice as many bloods as boys of their age in modern schools.

At 11 and 12 years, grammar school girls preferred comics to bloods. After that age, the reverse was true. Bloods and comics represented a smaller proportion of the total periodical reading of girls than of boys at grammar schools. This proportion decreased among the girls from 59 per cent. at 11 years to 6 per cent. at 15 years and over. Between 11 and 14 years, inclusive, grammar school girls read nearly twice as many bloods and about half as many comics as their opposite numbers in modern schools.

Newspapers.—96 per cent. of the grammar school children read one or more newspapers regularly. Each grammar school child named on an average 1·1 national daily papers, 0·5 London evening papers, 0·8 Sunday papers, and 0·1 local papers. The *Daily Mirror* read by 34 per cent. of the grammar school children was the most popular national daily, with the *Daily Express* (26 per cent.) second, and the *News Chronicle* (11 per cent.), third. The *Evening News* (27 per cent.) was the most widely read evening paper and the *News of the World* (19 per cent.) the most widely read Sunday paper.

Cinema.—The grammar school boys went to the cinema a little more than, and the girls a little less than, once a fortnight. 38 per cent. of these boys and 33 per cent. of these girls went once or more each week. Only 2·5 per cent. of the grammar school children did not go to the cinema. Modern school boys and girls went to the cinema nearly twice as often as grammar school children. The proportions of children who visited the cinema twice or more each week were more than three times as great among modern school boys and five times as great among modern school girls as were found among boys and girls of the same age group (11-14 years, inclusive) in grammar schools.

Radio.—Nearly 99 per cent. of the grammar school and 96·5 per cent. of the modern school children were radio listeners. Each grammar school child named, on an average 4·5, and each modern school child 3·3 programmes which were listened to with pleasure.

The most popular programmes among grammar school boys were Ignorance is Bliss (named by 54 per cent.), Twenty Questions (41 per cent.), Paul Temple (34 per cent.), and Dick Barton (33 per cent.). Among the grammar school girls, plays and serials (54 per cent.) came first, Twenty Questions (52 per cent.) second, Ignorance is Bliss (38 per cent.) third, and Light Orchestra (34 per cent.) fourth. There were marked changes in taste among boys and girls with the advance of years.

Children's Hour was more popular among grammar school than among modern school children. It was more popular with girls than with boys in each type of school.

Clubs and Associations.—Between 11 and 14 years, inclusive, 54 per cent. of the boys and 64 per cent. of the girls at grammar schools and 19 per cent. of the boys and 31 per cent. of the girls at modern schools belonged to some religious organisation. Among the boys at grammar schools, membership declined between 11 and 15+ years from 73 per cent. to 24 per cent. Among the grammar school girls, membership fluctuated between 59 per cent. and 68 per cent. during the same period.

Between 11 and 14 years, inclusive, 37 per cent. of the boys at grammar schools and 34 per cent. of boys at modern schools belonged to the Scouts or Boys' Brigade. 38 per cent. of the grammar school girls and 31 per cent. of the modern school girls belonged to the Guides or the Girls' Brigade.

30 per cent. of grammar school boys and 38 per cent. of modern school boys between 11 and 14 years, inclusive, belonged to cinema clubs. The corresponding figures among the girls were 20 per cent. and 26 per cent.

School Hobbies' Clubs (e.g., stamp, chess, railway, etc.), were mentioned frequently by grammar school boys. 31 per cent. of the 11 year olds belonged to one or other of these clubs. Grammar school girls and modern school boys and girls rarely mentioned such clubs.

School dramatic, music, scientific and other such societies were mentioned far more frequently by grammar school than by modern school children.

Favourite Activities Outside School.—The most popular activities among the grammar school boys were cricket (named by 61 per cent. of the boys), football (55 per cent.), cycling (37 per cent.), reading (36 per cent.), and swimming (35 per cent.). Among the girls, the most popular activities were swimming (66 per cent.), reading (65 per cent.), cycling (47 per cent.), tennis (38 per cent.), and walking (24 per cent.). Watching sport was included by 30 per cent. of the grammar school boys as a whole and by 43 per cent. of those of 15 years and over.

' Expressive ' activities such as acting and writing, were mentioned nearly four times as often, and non-participant entertainment (theatre, cinema, etc.), about twice as often by girls as by boys at grammar schools.

Tennis was mentioned four times as often by boys and six times as often by girls in grammar schools as by their opposite numbers in modern schools. Collecting was mentioned seven times as often by grammar school as by modern school girls. On the other hand, gardening and cinema were mentioned more often by both boys and girls and dancing by girls, in modern schools.

XI.—QUESTIONNAIRE.

1.—What is your sex ?
2.—How old are you ?
3.—How many older brothers and sisters have you ?
    How many younger brothers and sisters have you ?
4.—Name the things you like doing best outside school.
5.—(*a*) Name the comics and adventures (e.g., *Beano, Girls' Crystal, Champion,* etc.) you read regularly for pleasure.
    (*b*) Name any other magazines and papers (excluding newspapers) you read regularly.
    (*c*) What newspapers, if any, do you read regularly ?
6.—(*a*) Underline any of the following to which you belong :
        Scouts, Guides, Cinema Club, Boys' or Girls' Brigade, Church, Chapel, Sunday School, Pre-Service Unit, Youth Club, Social Club.
    (*b*) Name any other clubs or circles, not mentioned above, to which you belong.
7.—(*a*) How many times a week do you go to the cinema ?
        If you do not go weekly, how often do you go ?
    (*b*) What type of film do you enjoy most ? (e.g. Westerns, Musical, Cartoons, Crime, etc.
8.—Name the radio programme you listen to regularly with pleasure.

PSYCHOLOGICAL STUDIES OF THE MENTAL QUALITY OF THE POPULATION.[1]

By P. E. VERNON

(Professor of Educational Psychology, Institute of Education, University of London.)

I.—*The definition of quality.* II.—*Present population tendencies.* III.—*Evidence for and against a decline.* IV.—*Effects of environment on test scores.* V.—*Summary.*

I.—THE DEFINITION OF QUALITY.

In the recent report of the Royal Commission on Population,[2] the need was stressed for maintaining the quality as well as the size of the British population. But the definition of good or poor quality is a matter on which different people hold widely different views. We may disagree with Hitler's views, for example, but we have little scientific justification for any other definition that we happen to prefer. There is some truth in the gibe that people who aim to improve the quality of their national stock always try to impose their own characteristics; thus intellectuals want to breed a race of intellectuals. However, we are on fairly safe ground if we concentrate on general intelligence as measured by intelligence tests. For Spearman's underlying *g* factor of all-round ability is something that can be objectively defined and measured; it is not a matter of subjective opinions or ideals. Ability at every job in daily life and not merely ability for intellectual scholastic work depends on it to a greater or lesser extent. If we can maintain or improve the general intelligence level of the population we shall have, not only more children suitable for secondary and university education, but also more people fitted for skilled mechanical, clerical, and other such occupations. In the Fighting Services during the war, we found, for example, that this quality was essential for officers, N.C.O.'s, parachutists, and for specialists of every kind.[3] Moreover, people who score highly on intelligence tests usually show other characteristics that are generally considered desirable. They tend to be superior in physical health, and in moral qualities. This is proven by Terman's[4] well-known studies of gifted children, and by the fact that delinquents, criminals and members of what are sometimes called the social problem group, are usually below average in intelligence. Naturally, there are many exceptions to these trends, for the causation of socially desirable or undesirable behaviour is immensely complex, and intelligence is merely one factor. Psychologists would be the first to admit that there are many abilities and traits other than intelligence which might be considered important in a mentally and physically healthy and well-balanced population. But we know much less about these other traits. We cannot define nor measure most of them as objectively as we can intelligence. Also there are strong grounds for believing that other mental qualities are to a much greater extent than intelligence dependent on upbringing and education and less dependent on heredity.

---

[1] This paper was read as an introduction to a symposium held by Section J (Psychology) of the British Association for the Advancement of Science at its Newcastle meeting on September 5th, 1949. Other speakers were Prof. L. S. Penrose, Mr. J. Maxwell and Mr. D. Kennedy Fraser; the chair was taken by Prof. Sir Godfrey Thomson. An account of the whole meeting is to be published in *The Advancement of Science*, 1950.

[2] *Report of the Royal Commission on Population* (H.M. Stationery Office, 1949).

[3] See Vernon, P. E., and Parry, J. B.: *Personnel Selection in the British Forces.* (University of London Press, 1949).

[4] Terman, L. M., *et al.*: *Genetic Studies of Genius*, Vol. I (Stanford University Press, 1925).

II.—PRESENT POPULATION TENDENCIES.

It is generally known that the present birth-rate of the upper social and economic classes is much lower than that of the so-called lower classes, and that the former are, indeed, failing to reproduce themselves. For example, Burt[1] found that the average size of family among children attending one of best schools in a London Borough in 1920, was 2·9, whereas in one of the poorest schools it was 5·2. Another good illustration is given by Table I which is taken from the Merseyside Social Survey.[2] There, the present population was classified broadly into 30 per cent. non-manual, including professional, managerial and clerical; 40 per cent. skilled and semi-skilled, and 30 per cent. unskilled, and the size of families of each group of parents was found. From this it can be predicted that the next generation will amount only to 19 per cent. of children of the top group and to 44 per cent. of children of the lowest group. This phenomenon of the differential birth-rate has been characteristic of most western European countries, also of the United States, throughout the present century, and is causing serious concern to many students of population. Owing to our improved health services it is no longer compensated by a differential death-rate. Many more children from lower socio-economic families are not only born but also reach maturity.

TABLE I.

PERCENTAGE OCCUPATIONAL DISTRIBUTION OF EXISTING AND FUTURE POPULATION OF MERSEYSIDE.

	Non-Manual	Skilled and Semi-Skilled	Unskilled.
Heads of Families in existing population	30	40	30
Fathers of future population......	19	37	44

Social class is not, of course, the same thing as intelligence, but the same trend is apparent when we study intelligence as such. Thus Thomson[3] found that even within a single social class, namely coal miners, families containing the most intelligent children tended to be the smallest. To quote Burt:[4] " The conclusion to be stressed . . . is not so much that the smaller, well-to-do or professional classes are producing few children, but that among the far more numerous working classes it is still the most intelligent families who contribute fewest to the next generation."[5] Further investigations by Burt, Thomson and Cattell,[3] also Fraser Roberts's[6] comprehensive survey of the intelligence of children at Bath, and the Scottish Research Council's study,[7] all point the same way. Less intelligent parents tend to marry earlier, to bear children more rapidly, and to go on having them to a later age than do the more intelligent. Moreover these investigators agree very closely that the net result is likely to be a decline in the average intelligence level of about 2 points of Intelligence Quotient (i.e., 2 per cent.) per generation of about thirty years. Burt has cautiously estimated the figure at 1½ points per generation, and Table II is adapted from the figures that he has calculated.

---

[1] BURT, C. L.: *Intelligence and Fertility* (London: Eugenics Society, 1946). This memorandum gives a most useful survey of the whole field.

[2] See JONES, D. C.: " Eugenics Aspects of the Merseyside Survey " (*Eugen. Rev.*, XXVIII, 1936, pp. 103-113).

[3] THOMSON, G. H.: " The Trend of National Intelligence " (*Eugen. Rev.*, XXXVIII, 1946, pp. 9-18). During the symposium, Mr. Maxwell reported that much the same decline of intelligence with family size occurred among professional class children, and in other homogeneous social groups, as in the total Scottish population.

[4] BURT, *op. cit.*

[5] CATTELL, R. B.: *The Fight for Our National Intelligence* (London: King, 1937.)

[6] FRASER ROBERTS, J. A.: " Intelligence and Family Size " (*Eugen. Rev.*, XXX, 1939, pp. 237-247.)

[7] SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION: *The Trend of Scottish Intelligence* (University of London Press, 1949). See review in the present issue.

TABLE II.

INTELLIGENCE DISTRIBUTIONS OBTAINED IN 1920 AND PREDICTED AT LATER DATES ON THE BASIS OF FAMILY SIZES.

Intelligence Level.	I.Q.	Proportions in 1920	No. of children per family	Proportions expected in 1950	Proportions expected in 2000
Superior (Scholarship, etc.)	130 and over	1·8	2·3	1·4	0·8
Good	115-129	12·2	2·7	10·3	7·6
Average + ...	100-114	35·1	3·3	33·4	29·9
Average –	85-99	37·5	3·6	38·6	40·5
Dull and Backward	70-84	11·9	4·2	14·2	17·9
Very Dull (Feebleminded, etc.)	Below 70	1·5	4·7	2·1	3·3
Average I.Q.		100·0	—	98·5	96·0

The third column shows the percentages at each intelligence level in 1920 and the last two columns the proportions to be expected in 1950 and 2,000 if the most intelligent families continue to be too small. Perhaps even more alarming than the drop in the average from 100 to 98·5, then to 96·0, is the indication that the proportion of children of superior level would be halved, and the proportion at the very dull level more than doubled, by the end of the century.

III.—EVIDENCE FOR AND AGAINST A DECLINE.

Now though predictions based on sizes of families appear completely logical and watertight to the psychologist, the layman is apt to be suspicious and to ask for more direct proof of a decline in intelligence. Moreover, many professional geneticists agree with him, for they point out that we know very little about the mechanisms of human heredity.[1] It is actually conceivable that a population could maintain a stable composition in respect of intelligence in spite of greater fertility among its less than among its more intelligent members. To provide a direct proof of decline is, however, exceedingly difficult. Burt has carried out extensive testing with the various versions of the Binet, Stanford and Terman-Merrill scales between 1913 and 1939, and he claims that during this period there did occur a decline in average I.Q. level corresponding, not to 2 points, but to 1·3 points per generation. But he admits that he could not be certain of the representativeness or comparability of the samples he tested; also that any alteration is only too likely to be masked by the tendency of more intelligent families to migrate from some parts of the country to others. Much safer comparisons could be made if the complete population of a large area could be tested twice over at an interval of fifteen, twenty or more years—a population from which or into which little migration occurred. This has actually been done with all 11-year-old children in Scotland, and the results, as they stand, fail to show any decline at all between 1932 and 1947.[1]

---

[1] See PENROSE, L. S. : "The Galton Laboratory—Its Work and Aims" (*Eugen. Rev.*, XLI, 1949, pp. 17-27). At the symposium Professor Penrose described the genetic mechanism of heterosis which might compensate for any decline. He pointed out that both height and longevity are also negatively correlated with size of family, yet both of these are increasing rather than declining.

It may be seen from Table II that if a decline does occur, it is likely to be most noticeable at the extremes. Hence, even if we cannot detect an appreciable drop in the average we might still be able to prove an increase in the number of individuals of mentally defective level. Now two official surveys were carried out in 1907 and 1929, and these did indicate a serious rise in the proportion of defectives from 0·46 to 0·80 per cent. of the population. However, the 1929 Committee and its medical investigator, Dr. Lewis, concluded that most of the apparent increase was due to greater thoroughness of ascertainment. Many consider that the increase amounted to about one-third, or rather more than this in rural areas and rather less in urban areas. But others regard the figures as so inaccurate and incomplete that they do not prove any increase whatever.[2]

IV.—Effects of environment on test scores.

Even if we could overcome the difficulties of sampling or of complete ascertainment, there would still arise considerable doubt as to whether national intelligence is declining in the manner that the differential birth-rate would lead us to expect. The reason for this is that we do not know the precise extent to which intelligence as measured by our tests is an hereditary or genetic quality. If intelligence is partly or largely dependent on environmental factors, then any decline consequent on differential fertility might be offset by improved education, or by measures of social reform, or medical advances. The inability of the Scottish investigators to discover any loss over the past fifteen years might be due simply to improvements having occurred in these factors. One might even go further and suggest that the apparent decline due to differential fertility is itself an environmental phenomenon. The lower intelligence quotients of children in large families might arise, not from any genetic weakness in these families, but from the poorer circumstances in which such children are usually reared.

A study of the literature on heredity and environment reveals extreme divergences between the views held, even by psychologists and other social scientists.[3] The nature-nurture controversy arouses violent prejudices, and it is extremely difficult to arrive at unbiased conclusions. Often the same evidence is held by different writers to point in opposite directions. An attempt will be made to summarise this evidence very briefly, under four headings:

(1) *Parent-Offspring Resemblance.*—Many pedigree studies have demonstrated the tendency of children to resemble their parents in intelligence.

---

[1] Scottish Council for Research in Education, *op. cit.*

[2] Thus Professor Penrose pointed out that the figures show a far greater incidence at school age than either before 5 or after 14 years. This can only mean that ascertainment is easiest between 5 and 14.

[3] Cf. Schwesinger, G. C.: *Heredity and Environment* (New York: Macmillan, 1933). Also the XXVIIth and XXXIXth *Yearbooks of the National Society for the Study of Education*, 1928, 1940.

Thus the Galtons and Darwins illustrate outstanding ability running in families; the Jukes and Kallikaks illustrate the opposite extreme. Much data has been collected by standardised intelligence tests showing the higher intelligence level of children whose parents belong to professional or skilled occupations.[1] At the same time, as would be expected on genetic theory, the offspring do vary above and below the parental level. Only some 14 per cent. of children of mentally defective parents themselves fall within the mentally defective range. In fact, the amount of resemblance in intelligence is quite similar to that of physical qualities such as height which are known to be largely constitutional. None the less such data do not seem to prove anything at all, since parents normally rear their own children and naturally provide education for them, commensurate with their own. Thus the resemblance might equally well be attributed to environmental influence. Some studies, such as that of Lawrence,[2] have compared the intelligence of orphans with that of parents who have not reared them. Here there is still some correspondence, but it is approximately halved, suggesting, therefore, that child intelligence, as measured by the Stanford Binet test, is only partially inherited.

(2) *Studies of Twins and Siblings.*—The correspondence between the intelligence of pairs of identical twins, who, of course, have identical heredity, is virtually perfect, whereas the correspondence between pairs of non-identicals and between siblings (that is ordinary brothers or sisters) is lower, between first cousins lower still, and so on. This fits in with what would be expected on hereditary theory, and yet here also environment might provide an alternative explanation. Identical twins tend to be brought up far more alike than do siblings. Moreover the figures show a closer intellectual resemblance between non-identical twins than between siblings, and this should not occur since their genetic similarity is the same. Presumably, it can be attributed only to a closer similarity of environment.

Particularly valuable evidence is provided by the rare cases of identical twins brought up in different environments. Newman, Freeman and Holzinger[3] in America, have collected nineteen such pairs and show that, when the environments differ considerably in intellectual stimulus, quite considerable differences in I.Q. between the members of a pair do arise, ranging up to over 20 points. Shuttleworth[4] has carried out a very neat analysis of the scores of twins brought up together and apart, and of siblings, from which he concludes that a child's intelligence may be attributed:

64 per cent. to hereditary factors.
16 per cent. to environmental differences between families.
3 per cent. to differences in upbringing between children in the same family.
17 per cent. to joint heredity-environment factors, i.e., to the fact that families with good heredity usually also provide superior environment.

This conclusion is quite consonant with that reached by other approaches, and it brings out the point that one cannot really separate nature and nurture; they interact with one another from conception onwards.

---

[1] See NEFF, W. S.: "Socio-economic Status and Intelligence—A Critical Survey" (*Psychol. Bull.*, XXXV, 1938, pp. 727-757.). FLEMING, C. M.: "Socio-economic Level and Test Performance." (This *Journal*, XIII, 1943, pp. 74-82).

[2] LAWRENCE, E. M.: "An Investigation into the Relationship Between Intelligence and Environment" (*Brit. J. Psychol. Monogr. Suppl.*, No. 16, 1931).

[3] NEWMAN, H. H., FREEMAN, F. N., and HOLZINGER, K. J.: *Twins—A Study of Heredity and Environment* (Chicago University Press, 1937).

[4] SHUTTLEWORTH, F. K.: "The Nature Versus Nurture Problem" (*J. Educ. Psychol.*, XXVI, 1935, pp. 561-578, 655-681).

(3) *Evidence from Studies of Foster Children and Effects of Schooling.*—We realise now that the Intelligence Quotient is not an absolutely fixed or unalterable quantity. Although the majority of children when retested with a reliable test after one, two or more years, do get much the same result, to within about 5 points, some show much greater changes, especially over long periods of time. It has also been possible to prove that environmental factors are sometimes responsible for such alterations. Very interesting data is provided by foster children brought up by parents not their own. Freeman, Burks,[1] Leahy[2] and others in America have shown that those in better homes tend to show higher intelligence, or, in other words, that they come to resemble their foster parents. This does not, however, immediately prove environmental influence, since there always exists a tendency towards selective adoption. The more intelligent foster parents usually make more enquiries and try to adopt children of superior heredity. Burks[3] concludes that, when this factor is allowed for, good or poor foster-home environment can make a difference of about 10 to at most 20 points in the Intelligence Quotient, or, stated in another way, that the contribution of nature is at least twice and probably three times as great as that of nurture under the conditions normally met in our civilisation.

During the past ten years or so, however, a number of studies have been published from Iowa University[4] by Wellman, Skeels and others, claiming to show far greater effects upon the I.Q. Children whose mothers were said to be mentally defective, when sent to good foster homes, have risen to above average or superior intelligence, and Wellman literally advises foster parents who wish to adopt to disregard hereditary background. Conversely, groups of orphans who have been left in an intellectually depressing institution have dropped from an average I.Q. of 90 to mentally defective level. Again, children attending nursery schools have shown remarkable rises and these, it is stated, have persisted even up to the time when they became university students. The most recent study along these lines is that of Schmidt,[5] who claims to have turned 254 12-14 year children diagnosed as feeble-minded into a mentally, well-adjusted group, as a result of three years special schooling. Their mean I.Q. rose from 52 to 89, and a considerable proportion of them went on to a full high school course of study.

At the same time, it is fair to state that many, perhaps the majority, of psychologists in America[6] as well as in Britain, are suspicious of the Iowa, and of Schmidt's, methods and results, if only because such results run counter to all our experience of the trainability of mental defectives. The investigators do not seem to have established proper controls, and they were careless in their choice of tests, and in their statistical methods and calculations. Some of the opposition

[1] See XXVIIth *Yearbook of the National Society for the Study of Education*, 1928, pp. 103-316.

[2] LEAHY, A. M.: "Nature-Nurture and Intelligence" (*Genet. Psychol. Monogr.*, XVII, 1935, pp. 235-308).

[3] BURKS, B. S.: "On the Relative Contributions of Nature and Nurture to Average Group Differences in Intelligence." (*Proc. Nat. Acad. Sci.*, XXIV, 1938, pp. 276-282).

[4] See *University of Iowa Studies in Child Welfare*, XV-XVI, 1938-39; also XXXIXth *Yearbook of the National Society for the Study of Education*, 1940.

[5] SCHMIDT, B. G.: "Changes in Personal, Social and Intellectual Behavior of Children Originally Classified as Feebleminded" (*Psychol. Monogr.*, LX, No. 5, 1946).

[6] Cf. XXXIXth *Yearbook*. MCNEMAR, Q.: "A Critical Examination of the University of Iowa Studies of Environmental Influences Upon the I.Q." (*Psychol. Bull.*, XXXVII, 1940, pp. 63-92); KIRK, S. A.: "An Evaluation of the Study by Bernardine G. Schmidt." (*Psychol. Bull.*, XLV, 1948, pp. 321-333).

party state that their figures show no greater environmental effect than Burks or Shuttleworth found. It would be advisable, then, to suspend judgment until similar experiments have been repeated under more strictly controlled conditions.

(4) *Evidence from Factor Analysis of Intelligence Tests.*—Much uncertainty arises in investigations of heredity because many different intelligence tests are used, some of which are unreliable or poorly standardised. It is sometimes argued that the Stanford-Binet and other commonly used tests have considerable educational content, and that, therefore, environmental effects might be expected; but that intelligence itself is an inborn quality dependent only on properties of the Central Nervous System. This argument is, however, of doubtful validity, since we can recognise intelligence only in so far as it expresses itself. There seems to be little point in discussing a hypothetical innate quality unless we can measure it. It is preferable to define intelligence as the highest common factor or *g* that can be extracted from the best tests we can devise. It is important, therefore, to study what our tests do measure, and we have recently reached a much clearer understanding of this problem through the techniques of factor analysis.

Such analysis shows that the ordinary verbal group test as used, for example, for selection at 11+, is by no means a pure measure of *g*. The scores or I.Q.s depend to an almost equal extent on verbal and educational ability. Moreover, these scores are somewhat affected by practice at doing tests in general; that is, they measure first *g*, secondly verbal ability, and thirdly an ability to do well at psychological tests. The practice effect may make 5, or in the case of very unfamiliar tests, even 10 points, difference in I.Q.[1] There do, of course, exist non-verbal tests which are less affected by education, but they seem to be still more subject to practice effects and to other errors and weaknesses. The Stanford-Binet or Terman-Merrill individual scales probably give the best measure of general intelligence among children, and these show smaller practice effects than most group tests. Fortunately, then the Stanford-Binet has been the most widely used in heredity studies, and is the one which yields the most reliable results as to the respective contributions of nature and nurture that have been quoted above. Yet even this scale can claim to measure *g* only to about 50 per cent.; to about 25 per cent. it too depends on verbal and educational abilities, and the remaining 25 per cent. probably consists of various other minor abilities and errors.[2]

We must conclude then that psychologists have not yet solved the extremely difficult technical problems of measuring the general intelligence of population samples really reliably. The I.Q.s that are obtained at present are very impure mixtures, and far more research is needed into the construction of valid and practicable measuring instruments. Until this is done it seems premature to expect any very definite findings on problems of genetic change in the population.[3]

[1] Cf. VERNON and PARRY, *op cit.* Mr. Maxwell revealed that a considerable rise had occurred in Scottish children's group test scores between 1932 and 1947 in districts where tests are frequently employed; whereas there was scarcely any rise in districts where tests are seldom or never used. It should, however, be stressed that the Scottish investigators' conclusion as to the absence of any decline was based on Stanford-Binet and Terman-Merrill results, and not on the group tests themselves.

[2] Cf. VERNON, P. E.: *The Structure of Human Abilities.* In the press.

[3] Arising out of the discussion at this meeting, the present writer undertook to investigate the relation of numbers of siblings to factor scores, which would be derived from a battery of tests given to a large and representative sample of 18-year recruits. This should show how far *g*, as distinct from verbal-educational ability, is correlated with size of family.

V.—SUMMARY.

1.—Most definitions of good or poor qualities in a population are highly subjective, but the *g*-factor underlying intelligence test scores can be objectively defined and measured; thus it provides a useful basis for population studies.

2.—The lower average intelligence of large than of small families is well established, and this might be expected to lead to a drop of at least 1½ points of Intelligence Quotient per generation. Direct evidence of such a decline has been put forward, but is of doubtful validity, owing to difficulties of ascertainment and sampling. The complete populations tested fifteen years apart in the Scottish Mental Surveys showed no decline.

3.—This unexpected stability may be due to genetic mechanisms of which we are ignorant, but may also arise from environmental influences on test performance. Evidence on the nature-nurture controversy is outlined.

4.—Correlations between the intelligence of parents and offspring, or between pairs of twins or siblings, appear to conform to hereditary determination, but the lowered resemblance when twins are brought up apart or children are not reared by their own parents, points to considerable environmental influence. The investigations by several American psychologists claiming to show very large environmental modifications are regarded as insufficiently substantiated. It is concluded that nature normally contributes two to three times as much as nurture, though they cannot strictly be separated.

5.—The argument that intelligence is an innate quality, but that our tests are unable to measure it, is rejected. Both group verbal tests and versions of the Binet scale do have large verbal-educational content, and the former are particularly susceptible to practice or sophistication effects. Further research with tests that will give purer measures of *g* is needed before safe conclusions can be drawn about population changes.

MEASURING SUDANESE INTELLIGENCE

By G. C. SCOTT
(*Vice-Principal, Gordon Memorial College, Khartoum*).

I.—*Introduction: The first oral test.* II.—*The second oral test.* III.—*Results of oral tests.* IV.—*The first and second printed tests.* V.—*The third oral test.* VI.—*The third printed test.* VII.—*The Jenkins' non-verbal test.* VIII.—*Summary.*

I.—INTRODUCTION: THE FIRST ORAL TEST.

DURING the years 1946 to 1948, I carried out, under the auspices of the Gordon Memorial College, Khartoum, the first experiments in intelligence testing in the Sudan. The children with whom I was concerned were those in the Sudanese elementary and intermediate schools. A small and unknown percentage of the child population is admitted at the age of about seven or eight to the elementary schools, which give a four year vernacular course in the usual subjects. Some of these children are then selected by examination for the intermediate schools, where they are given a further four years' general education which includes English, the medium of instruction being still the vernacular. Thus Grades I to IV (mostly aged 7 to 12) are in the elementary schools, and from Grade IV children are selected for Grades V to VIII (mostly aged 11 to 16) in the intermediate schools. The demand for places in the schools varies a good deal, but is usually—even in the elementary schools—much in excess of the places available. About half the children have birth certificates, and the ages of the other half can only be guessed at. The great majority of them come from illiterate or semi-literate homes, and the language in which they are taught to read and write, "classical" Arabic, differs from the colloquial Arabic which they speak by more than as much as Chaucer's English differs from modern English.

I thought that in these circumstances a little-developed type of test—the orally administered group test—appeared the most suitable for first experiment; and I started with an adaptation of the Gabbani-Ballard Test.[1] This is an orally administered group test of British origin, but translated and adapted for Egyptian use. It consists of sixty-four scored items, and is administered in two periods of forty-five minutes each. The first period starts with five practice items of various types, and the second with one practice item. The scored items are not grouped together in sub-tests each of one type, but are mixed up. Every item is presented orally, while the necessary words or figures are written on the blackboard. Each child writes his responses. I changed the language of the test from colloquial Egyptian Arabic to colloquial Sudanese Arabic, and altered a few of the items which seemed unsuitable to the Sudanese children's environment. Then, thinking to standardize the test, I administered it to what seemed a representative sample of 900 birth-certificated children of ages between 10 and 16. The result was unpleasant. The correlation between age and test-score was microscopic in the intermediate schools and non-existent in the elementary schools.

No amount of bias in the sample could explain this result, and I began to understand the difficulties confronting anyone who tries to measure intelligence in a new country, especially one of the less "civilised" countries. First, the test items must be wholly suited to the children's environment; and, however well one knows the environment, it is impracticable to guess at the suitability; it must be discovered by experiment. The Gabbani-Ballard Test was, no doubt, suited to Egyptian children, but when it was used in the Sudan, even with prior adaptation, two-thirds of its items proved weak or wholly undiscriminative—some because they had become too easy, and everybody got them right; some because they had become too difficult, and everybody got them wrong; and many because of misunderstandings due to associations of ideas which differed in the different geographical and social environments. Secondly, the presentation of the items must also be suited to the environment.

---

[1] ISMAIL AL GABBANI, : Ikhtibar Al Zaka' Al Ibtida'i, Lijnat Al Ta'lif wa Al Tarjama, Cairo (Arabic); and P. B. BALLARD, *The New Examiner* (Hodder and Stoughton), Chap. XX.

An intelligence test does not necessarily measure all the intelligence a child has, but only all that he is willing and able at the time to show. The youngest and stupidest must understand what he is asked to do, even if he cannot do it; he must be interested in doing it, and trying his best to do it; and he must not be disturbed by irrelevant emotional factors, of which the most likely is being scared. Means to fulfil these conditions differ in different environments. What works in England may well not work in Egypt; and what works in Egypt may well not work in the Sudan. One can find the right way only by experiment. I, therefore, set out on a series of experiments designed to discover, (*a*) valid items, (*b*) efficient presentation.

II.—THE SECOND ORAL TEST.

(*a*) I tried out 400 items on 460 children, and selected from them eighty-eight items by the use of three criteria. They should show, first, not too great and not too small a percentage of correct responses, second, a percentage of correct responses significantly greater in any particular grade than in the grade below (where the children's mean age was a year less), and third, a good correlation between the tendency to respond correctly to them and the tendency to score a high total for all the selected items. I tried all the Gabbani-Ballard types, and a number of others, making twenty-one types in all, of which eleven survived. The ten *rejected types* were:

1.—" Best reason " type, e.g., " Why are donkeys useful animals? Because they live long? carry us? eat grass? have four legs?" This was one of the worst types.

2.—" Always " type, e.g., " Say what a man always has—cap, stick, head, house, clothes." These were either too easy, or failed to survive the other two criteria.

3.—Spatial visualisation, e.g., " A man went 10 miles west, then 10 south, then 10 east, then 5 north. How far was he from his starting point? " Correct response to this type showed no correlation with correct response to the others. (Yet it is a very good type in England.)

4.—Dictation of numbers. No correlation with school grade.

5.—Absurdities, e.g., " He knew it was a very old coin, because it was marked 60 B.C. Pick the best of four given reasons why the statement is absurd." No correlation with school grade.

6.—Catches, e.g., " I have a brother and two sisters. So has my brother. How many are we all together? " Good correlation with school grade, but failed to pass the consistency criterion.

7.—A type of classification—picking the word which includes the meanings of the other 4, e.g., " walk, dance, move, jump, run." I found it impossible to make this type sufficiently difficult without invalidating the items by including words affected by varying individual invironment.

8.—Word progressions, e.g., " Think of the proper order of these words, and then write down the middle one—sentence, page, book, word, letter." This type gave better results than those previously mentioned, but its consistency with other types was rather weak.

9 and 10.—Two types of number progressions, one of which was similar to type 8, while in the other the children were required to eliminate an extraneous number intruded into an easy arithmetical or geometrical progression. These types gave fair results, but not so good as those of the type of number progression adopted (No. 18 below)—perhaps because some intelligent children failed to grasp what was asked.

The eleven types of item which survived the experiments, and were eventually incorporated into a test, were as follows :

11 and 12.—Two types of word-classification, in one of which the child is required to pick out the one word whose meaning is different from the others, and in the other to pick out two words whose meanings are similar. Sixteen items survived, picked out of many duds.

13.—Rearranging the mixed up words of a sentence. This type was very hard to make, because the syntax of colloquial Arabic is not the same as that of classical Arabic, and yet there must be only one correct solution, whichever syntax one uses. But when this difficulty was surmounted, this type and type 14 produced very satisfactory figures. Six items.

14.—The same as 13, but instead of writing out the sentence, the child does what the sentence tells him (e.g., draws a circle with two dots in it). Four items.

15.—Filling in the missing word in a sentence. Five items.

16.—Opposites. Twelve items.

17.—Analogies, e.g., " Foot is to knee as hand is to . . . " Twelve items.

18.—Number progressions—adding two terms to an arithmetical or geometrical progression. Twelve items.

19.—Number classifications—like types 11 and 12, only with numbers instead of words. Four items.

20.—Instructions. I tried various types of instructions, and found that by far the most satisfactory was the alphabetical type used so much in the Moray House Tests : Eleven items. Also two items of the type " Write down the even numbers between 15 and 9."

21.—Items of the type, " If the day after to-morrow will be Sunday, what was the day before yesterday ? Four items.

I thus had eighty-eight items to make up into a test. The verbal bias was intentional, because the test's main practical use was to be for the selection of children for the intermediate and secondary schools.

(*b*) Faulty presentation not only reduced the score, but tended to invalidate the results, whether because some children did not understand what they were asked to do, or because they were uninterested, or afraid. It might be argued that only the less intelligent would fail to understand, so that the results would thus not be invalidated. But this is unsound, because failure to understand the requirement penalizes far more than inability to respond to it—there being many items all with the same requirement. I found that slight alterations of presentation could, sometimes, increase or decrease the correct responses by as much as 30 per cent.

I was, of course, very careful to follow Ismail Bey Al Gabbani in using colloquial and not classical Arabic, although the latter was the language which the children normally read and wrote ; and to take great care that the vocabulary should be within the compass of the youngest and least intelligent child. The aim was to test not knowledge of the meanings of words, which depends so much on environment, but the ability to relate ideas.

But I made, as the result of experiment, many alterations in the method of presentation. First, I introduced a " warming-up " period of forty-five minutes. I found that this was essential if the children were young, and beneficial if they were older. It provided the triple advantage of familiarising the children with the types of problem they would be asked to solve, banishing their fear of the strange white teacher with the foreign accent, and leaving them looking forward to what was to come—not as a test or examination, but as a jolly game.

Secondly, I prepared placards with the necessary words or figures on them, and so had no need to use the blackboard. This not only saved time, but increased the clarity and uniformity of presentation.

Thirdly, I collected all the items of each type together in sub-tests, and preceded each sub-test with an explanation, and with from one to three practice items. This again increased clarity, and together with the use of placards, saved so much time that I was able to present twelve practice items and eighty-eight scored items in the same time as the Gabbani-Ballard Test had taken for six practice items and sixty-four scored items, and yet include an increased amount of explanatory matter.

Fourthly, I experimented so as to get the best wording in the explanations, and the best order of items and of sub-tests.

Finally, I found that my own manner had also a considerable effect. The children would not show all the intelligence they had unless they were interested and unafraid ; and all of them would not be so unless my manner was friendly, cheerful, joking and encouraging. This leads to a dilemma. If different testers all attempt to be spontaneously friendly and cheerful, there will be a disastrous lack of uniformity in the presentation of the test. The only way out is for the method of presentation to be recorded in great detail and followed exactly by each tester, but in such a way that he appears to the children to be speaking spontaneously. This makes the oral type of test very difficult to administer ; but (as will be recorded), I found subsequently that printed verbal group-tests are useless for Sudanese elementary school children.

At the close of these preliminary experiments I was able to put together a test—the " Second Oral Test "—which promised better than the first one—the adapted Gabbani-Ballard. But at the start of standardisation I was faced with almost insuperable sampling difficulties. Even all the Sudanese schools together present only a small and biased sample of the total child population ; and the degree and perhaps also the nature of the bias vary very much between schools in accordance with varying local demand and social conditions. Neither do there exist in the Sudan statistics such as are available in Europe and America concerning the proportions in the community of rich and poor, townsman and countryman, and so on. Again, only some of the children in each school are available—those who have birth certificates. And there is reason to suppose that the birth-certificated children are not a representative sample even of their own school. For it is likely that they have higher intelligence than the average, since their parents tend to belong to the more successful classes of the community. I did not try, therefore, to obtain a representative sample of the total population, but made a provisional standardisation on 620 children in seventeen schools of differing types, hoping it was a fair sample of the birth-certificated school population. The children were aged from 9 to 15, and divided between Grades III to VIII in approximately equal proportions.

III.—RESULTS OF ORAL TESTS.

The results of this " Second Oral Test " were much better than those of the First. In spite of the negative correlation in any class between age and score (which proved as invariable in Sudanese schools as in English ones), there was a clear overall positive correlation. The raw score distributions for Grades III and IV, it is true, had marked positive skews ; and this could not be a true measurement, reflecting previous selection for intelligence, since the curves for Grades V to VIII, which were certainly more highly selected, all showed negatives skews. Both skews were clearly due to the age-range of the sample being bigger than the test could adequately deal with. It was too difficult for Grades III and IV, so that while the older and more intelligent children had scope to show their intelligence, there was a tendency to bunch together those of average intelligence and those below average. And the converse was true in the higher grades.

The factors of selection in the sample, which the figures brought out very clearly, made it impossible to standardise by taking as norms the mean scores of the different age-groups. Instead, I worked from the mean scores and mean ages of the various grades, on the assumption that in each grade the younger children with the higher intelligence would tend to cancel out the older children with the lower. This gave surprisingly plausible results, the means and distributions of intelligence quotients in the various schools turning out to be just what knowledge of the varying selective factors would lead one to suppose likely.

I made some rather abortive efforts to measure validity. I could not correlate my test with another known to be valid, because mine was the only one there was. There remained only correlation with teachers' estimates of intelligence and with examination results ; and neither was conclusive. The trouble about teachers' estimates was that both their validity and their reliability were very much in question. As to validity, none of the teachers was familiar with the technical meaning of " intelligence " ; and efforts to explain it were difficult, since none had any experience of intelligence testing, and since the Arabic language is deficient in modern technical terms. Some of the teachers certainly included common-sense, industry and attainments within the meaning of " intelligence," and some, perhaps, even obedience and good manners. As to reliability of the teachers' estimates : I made an effort to measure it by inter-correlating the independent estimates of four or five teachers. I did this with seven lots of children, obtaining forty-six correlations between pairs of estimates. The highest correlation coefficient was +0·88, and the lowest +0·21. The Mean and Standard Deviation of the values of r were, respectively, +0·64 and +0·254. After this I was not unduly discouraged when I found that the values of r for test-teacher correlations were much the same as for inter-teacher correlations. The highest was +0·90, and the lowest +0·20, with a mean of +0·6. The one case of +0·90 was obtained from a batch of twenty children, and the S.E. of r was 0·05. Correlating test-scores with marks in local school examinations, I found that r averaged only +0·54. But, perhaps one could not expect more, bearings in mind that the reliability coefficient of the examinations was unlikely to exceed +0·7, and that they were measuring many other variables besides intelligence.

As for the reliability of the test—it was given for a second time (after an interval of six months) to ninety-one children, and the scores correlated as follows :

Grades	S.D. of Scores		r	S.E. of r
	1st time.	2nd time		
III or IV	10·3	13·5[1]	+0·92	0·024
V, VI and VII	10·9	10·5	+0·88	0·042

---

[1] The increase in the S.D. is probably due to the fact that on the first occasion the test was too difficult for the younger children, thus causing the tendency to bunch, already noticed.

IV.—THE FIRST AND SECOND PRINTED TESTS.

While there were good reasons (the difference between the oral and written languages, and the illiterate environment) for the use of orally administered tests in the Sudan, they suffered from the great disadvantage of difficulty of administration. I, therefore, also made experiments with two tests of the more normal, printed type.

The " First Printed Test " was exactly in the form of a Moray House test, and its items were a selection, translated into Arabic, from six actual Moray House tests. The difference between spoken and written Arabic made the production of this test (and of other later tests) very difficult. The difference pervades vocabulary, grammar and syntax. If I wrote " classical " Arabic I was not using the children's mother-tongue, while if I wrote " colloquial " it would be the first time they had ever seen it on paper. I compromised by using, as far as possible, those words and those forms of grammar and syntax which were common to both forms of the language, or which, though not found in the colloquial, might be supposed to have become familiar to the mentally and educationally youngest of the testees. But this test, like the " First Oral Test," was given without preliminary experiment, and the figures it educed were even less satisfactory. Some items turned out to be much too easy, others much too difficult, and others to contain flaws which, though clear enough after experiment, could not have been guessed beforehand. The remaining items were both inadequate and unbalanced, and the presentation was not clear. As a test, it was useless. But it taught me, in addition to the necessity for experiment, three other important things. First, I found that the time required in the Sudan is longer than for a corresponding test in Great Britain. For instance, I administered this test to forty-two Gordon Memorial College students aged from 18 to 22, whose mean intelligence quotient was certainly a high one, and allowed them the forty-five minutes which is allowed in Great Britain. Only 20 per cent. of the students had time to attempt all the items, and 20 per cent. failed to get as far as item 77 (out of 100). Secondly, I found that the correlation between score and the time taken over the test was very low ; but of this I will give an account later. Thirdly, I found that a minority of children sometimes gave incorrect responses to a whole sub-test simply because they had not grasped what was required of them. This, of course, unduly penalised them.

The " Second Printed Test " was simply a printed version of the Second Oral Test, which I made in order to find out the effects of conversion from oral to visual presentation. Unfortunately, the content was not exactly the same, because the oral test included a larger proportion of " free response " as opposed to " multiple choice " items, and in order to simplify the scoring in the printed test, I had converted twenty-nine of the eighty-eight items to " multiple choice." This was a mistake, because it affected not only the items' degree of difficulty, but also (I think) the nature of the abilities required to solve them. An element of imagination enters into the solution of " free response " items that is lacking in that of " multiple choice."

When I correlated the " Second Printed Test " with the " Second Oral Test " in Grades V to VII, the coefficient was in the neighbourhood of +0·8. Bearing in mind that the reliability of the oral test was—in these Grades—less than +0·9, and that of the printed test certainly no more, and that the content of the two tests was not identical, I was encouraged by this figure to believe—what in fact proved possible later—that with improvement both in content and presentation a reliable printed test could be made for these grades. But in Grade IV it was a very different matter. Here the correlation between the two tests was only +0·4, and the mean score for the printed test was less than that of the oral by a whole year of Mental Age. It was clear both from study of the responses and from conversations with the testees that Sudanese children of these mental and educational ages find great difficulty in realising that the object of writing is to convey meaning—partly because they have not practised it for long, partly because the written language is not their mother-tongue, partly because of their illiterate home-environment, and partly, perhaps, because some teachers still tend to teach " classical "Arabic as material for correct recitation rather than as a vehicle for ideas. In the next grade above, these difficulties are surmountable. There the children are not only a year older chronologically. Mentally and educationally they are much more than a year older, because they are only a selected few from the grade below.

V.—THE THIRD ORAL TEST.

I did not offer the " Second Oral Test " for practical use, but based on it two further tests—the " Third Oral Test " for elementary schools, and the " Third Printed Test " for intermediate schools. The " Third Oral Test " was similar to the Second in content and form, with some improvements (learnt by experience) in the details of presentation. It consisted of thirty-four of the old items, and fifty-four new ones selected from 150 originals by the same criteria as before. This time I made no effort to find a representative sample even of the total school population, because I knew it could not be done, but tried only for a representative sample of birth-certificated children in the elementary schools of the capital. I administered the test to 672 such children in Grades III and IV in eleven schools, and also to 108 children in Grade V in two schools.

In order to standardise, I determined, as before, to work from the mean age and mean score of the Grade. But I wanted some confirmation that this was a reasonable thing to do. I had, therefore, given the test twice, at an interval of seven months, to 160 children, to find out the rate of increase in their score. The mean increase was 12 points, with the Standard Error of the Difference equal to 0·4. I also found the mean increases in score for the three separate age-groups of 9, 10 and 11 to be 12·9, 12·5 and 11·3, respectively, with the Standard Errors of the Differences 0·9, 0·9, and 1·1.

How much of this score-increase of about 12 points in seven months was due to practice effect, and how much to increased mental age? Observation of the children suggested, contrary to experience elsewhere, that practice effect was negligible. Before the second testing they were told they were going to have " another test, like the one you did before." No child showed any sign of realising that any items were common to the " two " tests, although leading questions were asked immediately afterwards. They had quite clearly forgotten all the items in the interval; and the same old explanations and jokes seemed to come quite freshly. Moreover, the absence of significant practice effect was confirmed by the fact that the rate of increase of 12 points in seven months, or 20 to 21 points per year, was exactly that indicated by the difference between the mean scores at a single testing of those in Grade III (mean age 10·4) and those in Grade IV (mean age 11·5). This difference, which could not possibly be influenced by practice effect, was 22 points at the first testing and 23 points at the second, for a thirteen months' difference in age, or 20 to 21 points, as before, for a year's difference.

I thus had two fixed norms, one for the age of 10·4 and the other for 11·4, together with evidence that the rate of increase of mental age corresponded to the same 20 points per year between the ages of 9·7 and 12·0. Thus all the norms between those ages could be fixed. I had, as yet, however, no norms above the age of 12·0; and it was unlikely that the rate of increase in score would remain constant, since the norm for age 12·0 was 57 points, and the maximum possible score 88 points. I made three guesses at the shape of the curve connecting score with mental age above the age of 12, allotted I.Q.s by each of them to all the children of 12 and upwards who had taken the test, and adopted the curve which made the mean I.Q. by the second testing most nearly equal to the mean I.Q. by the first testing. The norms so fixed cannot have been wildly inaccurate since the I.Q.s of the 672 children derived from them averaged 100·8 in Grade III and 101·1 in Grade IV. The Standard Errors were 0·6 and 0·7, respectively, thus these means did not differ significantly from each other or from 100.

I made, as with the "Second Oral Test," some attempt to test validity by correlating test scores with teachers' estimates of intelligence. The experiment was made with 40 Grade V children in one class, whose intelligence was assessed independently by three teachers. The correlations of each set of estimates with the other two combined were +0·71 (S.E. 0·078), +0·80 (S.E. 0·056) and +0·68 (S.E. 0·084). All three combined together correlated +0·62 (S.E. 0·098) with the test scores. This was inconclusive, of course, owing to the large Standard Errors; but, bearing in mind the small Standard Deviation of the test scores, it was not discouraging. On these figures, a Standard Deviation of 20 for the test scores (corresponding to a year of mental age) would give[1] a correlation of +0·85.

I also calculated two correlations between the test scores and marks awarded by two different examiners in the Intermediate School Entrance Examinations, which took place 8 to 9 months after testing. In the first case (thirty-four pupils from two schools) I got the surprisingly high figures of $r = +0{\cdot}82$, with S.E. 0·057. In the other (seventy-three pupils from four schools) $r = +0{\cdot}61$, with S.E. 0·073. In assessing this figure, it may be noted that ability in teaching and coaching has a very great effect on the marks obtained in the examination; for the mean marks of pupils in yet another school, famous for its efficiency exceeded by 40 (maximum 200) the means of the four (S.E. of the Difference= 3·2), although the markers were the same, and the differences between the mean mental ages of the five schools were barely significant.

The reliability of the test was measured by correlating first the scores of the 160 children who had taken it twice, and then the I.Q.s derived from them. In each case $r = +0{\cdot}94$, with S.E. 0·011. The S.D. of the scores was 15·4 points,[2] of the I.Q.s 10·7 points, of the mental ages 9·2 months, and of the chronological ages 12·1 months. I then picked from the 160 children a sample of sixty-seven in such a way that the I.Q.s were distributed normally with a S.D. of 13, and got r I.Q.=0·97, S.E. 0·008. The "Third Oral Test," then, appeared to be sufficiently sound for use in Sudanese elementary schools.

VI.—THE THIRD PRINTED TEST.

The "Third Printed Test," on which I was working simultaneously with the "Third Oral Test," was for intermediate schools. I used for it sixteen items from the "Second Oral Test," which had not been included in the "Third Oral Test," and seventy-three new items selected from 230 originals by the same criteria as before. With a view to reducing the amount of explanatory and

---

[1] By the formula $R = \dfrac{\frac{\Sigma}{\sigma}\,r}{\sqrt{1 + \left(\frac{\Sigma^2}{\sigma^2} - 1\right) r^2}}$

[2] Therefore, the "Standard Error of the Estimated Score" was 5·7 points, corresponding to about 4 months of mental age.

$$\sigma_{\text{est.}} = \sigma\sqrt{1 - r^2}.$$

practice matter, and the danger of children failing to grasp what was required of them, I used only eight instead of eleven types. Seven were types employed previously in the oral tests (Nos. 11, 13, 15 to 18, and 20 on p. 3). The eighth was a type of reasoning item from the Moray House tests.

In devising the presentation, my previous experiments with the First and Second Oral Tests proved very useful. For instance, I knew that more "warming-up," more explanations, examples and practice items were necessary than in Great Britain, and I had a better idea of the type of Arabic which would be comprehensible. The most difficult question was that of timing. Because I had under-estimated the time required for the "First Printed Test," a large proportion of the children had failed to attempt all the items. This enabled me to make a correlation between the number of items attempted and the score achieved, and I found r to be less than +0·2, with a Standard Error of 0·09. There was, indeed, a noticeable tendency for high scores to have attempted many items; but there was none for those who attempted many items to score high. Since this indicated that in the Sudan and with this type of test the correlation between fast working and correct working was probably low, I adopted a very generous time-limit, which allowed all but a very small minority of hopeless cases to finish the test. And I added at the end a number of unscored "time-wasting" items to keep those who finished early from disturbing the others.

A very provisional standardisation was made, using the same methods as for the "Third Oral Test," on 572 children approximately equally divided between Grades V, VI, VII, and VIII. In addition, the test was administered to thirty-seven pupils in Grade IX (secondary school) for comparison, and to eighty more in Grade V who had previously taken the "Third Oral Test," for the purpose of tying the Printed Test's norms to those of the Oral Test. Considering that the Printed Test, being designed mainly for the measurement of high mental ages in Grade VIII, was much too difficult for Grade V (the mean score was only 21 out of a possible 89), also that the S.D. of the I.Q.s of both tests for these eighty children was only 9, it was not discouraging to find a correlation between I.Q.s from the two tests of +0·87, S.E. 0·027.

As for validity, I correlated the test scores of 236 boys in Grade VIII with their marks in the Final Intermediate School Examinations (in the subjects of Arabic, English, Arithmetic and Geography, weighted 27, 40, 20 and 13, respectively), held about two months later, and got r=+0·66, S.E. 0·037. (The S.D. of the test scores was 15·5.) I did not, this time, correlate the test scores with teachers' estimates, as I had become doubtful of their value. But four correlations made earlier between teachers' estimates and the scores on the preliminary experimental oral version of the test were +0·60, and +0·61 in Grade VII, and +0·88 and +0·90 in Grade VIII.

Reliability was fair, but not so good as that of the "Third Oral Test." The correlation for 171 children who took the test a second time after an interval of one month was +0·93, S.E. +0·009, with S.D. of the scores =17·7. For 182 other children who took the test a second time after an interval of six months, r=+0·91, S.E. 0·012, with S.D. of the scores =17·4[1]. Practice effect was noticeable, even after six months, in contrast to its absence in similar circumstances with the "Third Oral Test." The explanation may be (i) the effect is greater with older children; (ii) the oral medium being more familiar than the visual, and the "warming-up" of the oral test being more efficient than that of the printed one, the material on the first occasion is less strange with the oral test, so that there is less room for practice effect to show itself.

---

[1] Taking r as 0·92 and S.D. as 17·5, the "Standard Error of the Estimated Score" works out at 6·9 points, corresponding to 6·4 months of Mental Age, comparable with 4 months for the oral test.

An encouraging feature of the " Third Printed Test " was that in examining the responses we found very few of the signs—so noticeable in the First and Second Printed Tests—of failure on the part of the children to understand the requirements. By an increased amount of explanatory and practice matter, and by very great care over the wording and lay-out, we had made a fair success of making " classical " Arabic intelligible to Sudanese children. In this I was very greatly helped by a Sudanese colleague, Mr. Derar Saleh, who had joined me while this test was still in preparation.

There is also a " Fourth Printed Test," intended as an improved alternative to the Third. It is on the same general lines, and has been made by my colleague by means of experiments of the kind already described. But it contains a number of improvements (both of content and presentation) which were suggested by examination of the responses to the " Third Printed Test." One of the improvements is the allotment of a few points to the "time-wasting " items at the end of the test. It appears that with the generous time-limit given and with the low correlation between fast working and high scoring, the careful plodder—owing to his greater care—tends to score slightly higher than the careless dasher of equal mental age. The extra points at the end may enable the dasher to make up his score to what it should be. The " Fourth Printed Test " is at the time of writing only just clear of the press, so that no figures are yet available concerning it.

VII.—THE JENKINS NON-VERBAL TEST.

I also spent a great deal of time in experimenting with a non-verbal test. Individual non-verbal tests, and, therefore, tests that made use of apparatus other than printed matter, were impossible. Thus I was left with the choice between pictorial tests and those using non-representational diagrams and patterns. I studied several pictorial tests. Among them was R. A. C. Oliver's " General Intelligence Test for Africans."[1] Though this consisted of 164 items, a reliability coefficient of only +0·81 was claimed[2] and it appeared to have been little used and not standardised. I examined other pictorial tests, noticeably the Otis and Cattell,[3] but it was obvious that however good they might be in Great Britain, they were quite unsuited to Sudanese children, owing to the foreign nature of the subject matter of the pictures. As I had not the means to make a pictorial test of my own, I decided to use diagrams, and chose J. W. Jenkins's " Scale of Non-Verbal Mental Ability."[4] We had to make a special version of it, not only because the instructions and numbering had to be in Arabic instead of English, but because—since Arabic is read from right to left—some of the diagrams, for example the progressions and analogies, had also to be reversed. We took very great trouble over the translation of the instructions, well knowing the high degree of simplicity and clarity necessary for comprehension by children in Grade IV.

But the very first experiment was disconcerting. I administered the test to fifty-two boys (mean age 15·11) who had just been admitted to Grade IX (secondary school), and must have had (owing to two selections) a mean I.Q. of at least 110. I also tested them by the oral version of the " Third Printed Test "

---

[1] The Government Press, Nairobi.

[2] Page 3 of the *Manual of Directions*.

[3] Both George G. Harrap and Co.

[4] The National Foundation for Educational Research in England and Wales.

(which was not then ready in its final, printed, form). Their mean score in the Jenkins' Test (out of a maximum of 85) was 58, which is the score for British children of I.Q. 111 and age 12·11 ; and the correlation coefficient between the two tests was only +0·3, S.E. 0·12. (The verbal test scores had a S.D. of 12·6, and the Jenkins Test scores 10·0.) Yet the correlation in Great Britain between the Jenkins Test and the Cotswold and Moray House Tests is quoted as +0·81. (The S.D. of the sample is not quoted. Some authorities consider that correlations between verbal and non-verbal tests are more likely to approximate to +0·6).

I then started out on a long series of experiments to find out whether alterations in the presentation of the Jenkins Test would increase correlation with my verbal tests. I introduced an oral " warming-up " period with fifteen practice items displayed on placards. In the test itself I not only read aloud the instructions (which were, of course, in simple " classical " Arabic), but repeated the sense of them in colloquial Arabic, holding up the booklet and pointing as required. And remembering my experience with verbal tests I increased the time allowed to a period which enabled all but the very slowest to finish (this meant only an increase of five minutes in thirty).

The results of these changes were remarkable. I now got a mean score in Grade IV (mean age 12·0) only 4 points below the original mean score in Grade IX (mean age 15·11, and twice selected by examination). And I found that the mean mental age of the children as shown by the British norms of the Jenkins Test was now about the same as that given by the norms of the "Third Oral Test " instead of being about two or three years below it, as the first experiment had indicated. But the correlation between the verbal tests and the non-verbal test in its new presentation, although higher than before, remained low. I gave the Jenkins Test and the " Third Oral Test " to 100 children in Grade IV and 100 in Grade V ; and on correlating the scores found :

Grade	r.	S.E. of r.	S.D. of Verbal Test.
IV	+0·58	0·07	13
V	+0·38	0·09	12

I also gave the Jenkins Test orally and individually to forty children who had already taken both the " Third Oral Test " and the Jenkins Test as a group test. All their scores increased considerably, but the correlation between the Jenkins Test and the verbal test remained no greater than before.

I also measured the reliability of the Jenkins Test by giving it twice to 100 children in Grade IV and 100 in Grade V, and on correlating the two sets of scores, got—not the r=+0·95, which is stated as the British figure, but :

Grade.	r.	S.E. of r.	S.D. of Scores.	
			1st Time.	2nd Time.
IV	+0·82	0·032	10·3	11·3
V	+0·68	0·053	8·8	8·4

It appeared, therefore, that in the Sudan the Jenkins Test was of little use in forecasting educability, and I abandoned it. I had not the time to experiment with other non-verbal tests.

Considerations of space prevent me from acknowledging all my very great obligations; but I cannot pass over my debt to Ismail Bey Al Gabbani, Moray House, the National Foundation for Educational Research in England and Wales, and above all to Professor P. E. Vernon who, while not responsible for my errors, has throughout my work generously given me the criticism and advice without which I could have done nothing.

VIII.—SUMMARY.

An account is given of the first experiments in intelligence testing to be undertaken in the Sudan. It was found that translations of foreign verbal group tests with *a priori* adaptations were invalid and unreliable. Both content and presentation were faulty. Local tests, based on extensive experiment, were, therefore made, and these turned out to be apparently valid and certainly reliable. While group tests could be administered in the usual printed form to the older children, it was found necessary to develop P. B. Ballard's oral technique of administration with the younger children.

Among the difficulties to be surmounted were the general illiteracy of the population, and the fact that the written and spoken languages are not the same. It was found that more preparation of the required emotional atmosphere, more explanation, more practice, and more time for the test, were needed than in Great Britain. Sampling difficulties precluded more than limited and provisional standardisations.

A non-verbal, diagrammatic, group test was also tried, but after extensive experiment was abandoned as being of little use in the Sudan for forecasting educability.

CRITICAL NOTICE.

THE TREND OF SCOTTISH INTELLIGENCE.

The Scottish Council for Research in Education (Univ. of Lond. Press., pp. xxviii +151, 7s. 6d. net).

As readers of this *Journal* will already know, the Scottish Council for Research in Education has recently completed what is perhaps the most ambitious of its many valuable investigations in the field of educational psychology. It has carried out a second psychological survey of a complete age-group, comprising all the accessible children aged eleven in the country, on lines somewhat similar to the survey undertaken in 1932, but with the addition of an informative ' sociological questionnaire.' A preliminary report on the inquiry has now been published ; and will prove a mine of information to subsequent investigators. It consists of nine chapters by eight different writers. Together they provide a detailed account of the arrangements made for the group and individual testing, and a general review of the main results.

As the preface explains, the primary purpose of the inquiry was to obtain " light on the causes of a remarkable fact, namely, that the results of intelligence tests show that the average score of members of large families is less than that of members of small families." Both in the daily press and in educational periodicals, there seems to have been some misunderstanding of the assumption that was actually being investigated. A number of popular writers have apparently supposed that the advocates of such inquiries cherish a belief that the mere fact of being a member of a large family is liable to handicap or hinder the young child's mental growth, and that " the only child, or the child of the small family, is better looked after intellectually as well as physically." That is not the issue here. Accordingly, it may be helpful at the outset to state exactly what the problem is, and how it has arisen.

At the close of the last century, there was, in certain educational quarters, a widespread alarm lest the innate abilities of the nation were deteriorating.[1] The need to relax the ' standards ' suggested in the original Board of Education code had suggested to the minds of many school inspectors that the level of educational attainments was declining. Soon afterwards, in a series of memoirs with the somewhat grim title, ' Studies in National Deterioration,' Karl Pearson and his collaborators showed that the birthrate in the poorest London boroughs was nearly twice as high as in the well-to-do, and that these differences had been increasing during the past thirty years. They argued from this that the least efficient members of the community must be multiplying at the highest rate. With this problem in view, a Committee of the Anthropological Section of the British Association urged the importance of a mental and physical survey of the general population, and drew up a tentative scheme. And when, in 1913, owing largely to the advocacy of Galton and Sully, the London County Council appointed a psychologist, it was agreed that one of his main tasks should be to undertake periodic surveys of the school population with standardized mental and scholastic tests.

---

[1] Much of the interest in this problem was due to Galton's work during the last half of the nineteenth century and to the efforts of the Eugenics Society (as it was originally called), founded under the presidency of Galton early in the present century. Half the expenses of the Scottish Council's research have, in fact, been generously defrayed by this Society. It may be added that Galton was fully aware that " poverty is not necessarily a measure of inefficiency " ; and it was primarily for this reason that, long before the days of Binet, he had endeavoured to devise intelligence tests. The construction of psychological tests for the anthropological committee was subsequently put in the hands of McDougall, and it was with this object mainly in mind that his earliest disciples (like Spearman and myself) undertook our first investigations in the schools near Oxford. The critics who object that " it is unfair to make wholesale assessments of normal school children by devices originally intended merely to help medical officers grade the feeble-minded " are thus inverting the true history of mental testing. Mental tests were originally devised for wholesale surveys ; it was much later that they were put forward for testing defectives.'

One of the earliest results was the discovery that the birthrate among families from which the duller children were drawn was about twice as high as among those supplying the brighter junior county scholars. What seemed most significant of all was the fact that, even within the same social and economic class, fully significant correlations were found. From the correlations observed (about 0·20) it was calculated that, if no counteracting factors were operative, the intelligence quotient would drop at the rate of about 1½ points per generation; and (what would be far easier to verify) " in little over fifty years the number of pupils of scholarship ability would be approximately halved, and the number of feebleminded almost doubled." Dr. Cattell, in two rural surveys, also financed by the Eugenics Society, obtained still more alarming figures; and declared that the nation seemed heading for " a galloping plunge towards intellectual bankruptcy."[1]

But, as was pointed out in my reports at the time, " an armchair calculation is no safe substitute for direct evidence." And here comes the most intriguing feature of these inquiries. My own opinion was that " changes so great as those deduced could hardly fail to excite spontaneous comments from teachers, medical officers, and officials of the Education Department." On repeating the surveys in London, the test-results, it is true, seemed to point to a small increase in the number of mental defectives; and the marks of the junior county scholars seemed to indicate a decline in mechanical arithmetic and spelling. As to intelligence, the evidence seemed to indicate the probability of a small decline—smaller than the calculations suggested, but large enough to be disquieting. However, " the only conclusion established beyond reasonable doubt was that a carefully planned inquiry, on a large scale in several different areas, was urgently needed."

When the Royal Commission on Population was set up, at the close of the war, it naturally gave special attention to this problem It began by inviting memoranda on the subject from psychologists and geneticists. The evidence was still conflicting; and in 1945 Sir Alexander Carr-Saunders (Chairman of the Population Investigation Committee) approached the Scottish Research Council with the request that it should repeat the survey that it had already carried out some years before.

As in the London surveys, the Scottish inquiry was based (1) upon a broad investigation carried out by group tests, supplemented (2) by a more intensive study by individual tests. In London, it was thought wiser not to use the same *group* test in successive surveys, but to keep the *individual* test the same for purposes of standardization. In the Scottish inquiry, on the other hand, the same group test was used on both occasions, but the individual test was changed. Unfortunately, in the 1932 survey, no data were procured in regard to size of families. But in the 1947 survey, the negative correlation found by many previous investigators was fully confirmed. With the results of the group test the correlation was −0·28; with the results of the individual test, −0·32.

The main interest lay, however, in a comparison between the average marks for intelligence obtained in 1947 and those obtained fifteen years earlier. What, then, are the figures? With the group test, which had a maximum of 76 marks, the average score of the Scottish children rose from 34·5 in 1932 to 36·7 in 1947. With the individual test the average intelligence quotient rose from 101·6 to 102·5. Thus the child of to-day is not duller but brighter than his predecessor of fifteen years ago. The figures have evidently come as a surprise to the investigators; and in various parts of the report several considerations are tentatively put forward as offering possible explanations for the result. However, not all the possibilities seem to have been explored; and, as future investigators will no doubt wish to repeat such surveys, it may be worth while to attempt a systematic summary of the more obvious causes, since some of the most troublesome disturbances might conceivably be avoided by suitable modifications in the plans. Similar factors were also noted in the course of the London surveys; and there it proved by no means easy either to verify their influence or to eliminate it.

---

[1] *The Fight for our National Intelligence*, 1937.

1.—The tests employed were originally devised in the hope of measuring *innate* intelligence. Hence the first hypothesis to be considered is that the innate intelligence of the Scottish nation has actually improved. This explanation has apparently struck the investigators as so improbable that they do not mention it in their list. Yet there is a widespread popular notion that the inborn abilities of nations may rise as well as fall; historians are fond of citing Athens in the age of Pericles, England in the age of Elizabeth, and Europe during 1600 to 1700—" the century of genius." No doubt, in the fuller discussions we are promised later on, the investigators will give reasons for finally rejecting this somewhat optimistic speculation.

2.—Every psychologist would agree that even the best intelligence test cannot profess to be anything like a pure and perfect test of sheer innate ability. The results of such tests, particularly when obtained by different investigators using a group procedure, and when left unchecked and unadjusted, are apt to be affected by various *environmental* influences.

(*a*) First, performance in group tests of certain types may undoubtedly be affected by improvements in health and physical fitness, which bring with them increased mental alertness. Where wages are low, it is the member of the large and overcrowded household who is likely to suffer most, both physically and culturally. Now we are frequently assured that during the last decade or so the *health and home conditions of the poorer children* have been steadily improving; and it may be remembered that the five or six years which preceded the 1932 survey were years of economic depression. It is by no means inconceivable that such an improvement might add a point or two to the equivalent I.Q.

(*b*) During the same period both teachers and children, especially I fancy in Scotland, have grown more and more familiar with the devices for intelligence-testing. When intelligence-tests were first employed, their novelty proved a little alarming to the young examinees, particularly those who were not particularly stable or adaptable; and even a mild touch of bewilderment is apt to reduce efficiency. Nowadays, however, by the age of eleven, a large proportion of pupils are pretty sure to have undergone some kind of intelligence test. During recent years, too, ' new-type ' techniques have been used in examinations of every kind, and have even been exploited as aids to teaching in the classroom and the textbook. The experience and training thus obtained, together with the resulting ' *test sophistication* ' (as it is somewhat inappropriately called), may have considerably raised the marks secured in the more recent survey.

There is, indeed, some factual evidence corroborating this suggestion. At the recent meeting of the British Association, Mr. Maxwell reported to the Psychological Section the results of a more recent analysis of the Scottish data

In those areas where intelligence testing is a routine procedure, the 1947 pupils showed a rise of 3·2 points over the score obtained by their predecessors in 1932; on the other hand, pupils who had been subjected to no such tests, at least during the preceding school session, showed a rise of only 0·4 points.[1]

(*c*) In London, whenever it became known that our test-surveys were to be repeated, one of our greatest difficulties was to avoid the effect of deliberate *coaching* in the tests it was assumed we might use. This preliminary preparation was quite honest and open. Teachers would argue that " it was not fair to the children to expect them to sit for an examination without letting them know the kind of questions they would be asked"; and many would add that it was not fair to the teachers either.[2] In Scotland, we are told, when it was decided to use the 1932 test for the 1947 survey, the test was withdrawn from circulation, and " great care was taken to ensure that it was not generally known that the 1932 test was to be used." But, as a Scottish headmaster tells me, he for one guessed that *something like* the previous test would probably be adopted, and knew it was to be found printed in full in the 1932 report. When he found that *precisely* the same test was used, he " felt decidedly guilty." It was in the hope of evading these difficulties that in London we avoided using precisely the same group test for two successive surveys; and it is instructive to observe that Dr. Fraser-Roberts himself, in his section on 'planning future surveys,' concludes that " it is quite unnecessary to use the same group test for different populations; only the individual test need be the same."

3.—The difference in the results of the individual tests is fully discussed in a chapter by Dr. Fraser Roberts. In the 1932 survey the test employed was the Stanford Revision of the Binet scale. Here there was a *change of tests*. In the 1947 survey, however, "for a variety of reasons" the Committee decided to substitute the L Form of the Terman-Merrill Revision. Thirty-nine boys and fifty girls were tested with both; and from their results it appeared that the Terman-Merrill revision was slightly easier than the Stanford. Accordingly, a regression equation was formed for converting Terman-Merrill I.Q.s to Stanford I.Q.s. When this correction was applied, it appeared that the average I.Q. for the boys rose from 99·9 in 1932, to only 100·5 in 1947, while the average I.Q. for the girls dropped from 98·6 to 97·9. Both changes are statistically non-significant; so that here, it would seem, there is no evidence that intelligence has either deteriorated or improved. The only criticism that I would venture to make is that the number used for obtaining the regression equation was regrettably small.

---

[1] *Times Educational Supplement*, September 9th, 1949, p. 618. The Report itself states that directors of education were requested to postpone their customary use of intelligence tests for promotion purposes until after the June survey, but that many directors " could not accede to this request." In 1947 there must have been far more directors using such tests than in 1932; and consequently this may have had a considerable influence.

[2] Many years ago, when I myself sat for a College scholarship, I found that one of the papers regularly included a 'test of general knowledge'; and, with the aid of my form master, studiously crammed up likely topics. I discovered " who fished the murex up, and what porridge had John Keats," and had the luck to be asked that question. I subsequently learnt that the college examiner was horrified at my immorality in " seeking to circumvent the purpose of the paper."

To-day, most teachers realize that an intelligence test is intended to assess the child's own innate ability, not the teacher's skill in teaching. Nevertheless, a few still seem unaware of the difference. Many often fear that some reflection will be passed on their school or on their own efficiency if the pupils under their care do not get such high marks as their predecessors under a former head teacher or as another school in the neighbourhood. Recently, when I was concerned with setting annual intelligence tests for entrants to a certain college, teachers continually wrote asking for a 'specimen examination paper' or a 'syllabus of the topics likely to be asked.'

4.—One factor which might conceivably influence the apparent trend of intelligence, and which, I fancy, is not mentioned by the investigators, is that of *migration*. In the London surveys, this appeared to be by far the commonest and most important cause of changes in the average mental ratio (or I.Q.) for a given locality. Now, at the beginning of the present century, the way in which abler Scotsmen migrated South had become almost proverbial. But, since the economic depression after the first world war, this movement South, so at least it has been alleged, has appreciably diminished, while a good many skilled English workmen and administrators tended, for a time at any rate, to cross north of the Tweed. It would be interesting to learn what evidence, if any, there may be for such changes, and whether it can have affected the results.

5.—I myself am convinced that educational surveys should, if possible, aim at covering *the whole age-range* of the junior school, at the very least. In the largest of the earlier London surveys, we endeavoured to get representative samples for all age-groups ; in the later, we generally planned a complete survey of a single age-group (nearly always the age for the junior county scholarship examination). But it became clear that with the latter procedure we were risking an undesirable form of selection. With only a single age group the boys and girls representing the largest families would include fewer first-born children than those representing the smaller families ; and, if the average intelligence of the first-born is (as some have maintained) slightly different from that of the later born, this of itself introduces a possible complication. Moreover, it is further conceivable that fifteen or twenty-five years ago the later-born children in an over-populated home may have suffered more from overcrowding and overfeeding, and that to-day an improvement in the standard of living may have partly removed such handicaps. Professor Thomson deals very fully with these possibilities in his admirable chapter on ' intelligence and family size.' On the whole, he concludes that such effects are not likely to be of great importance in the Scottish surveys. But they are certainly considerations to be borne in mind in future plans.

6.—During the years covered by my work in London schools, the inclination to limit the size of families spread pretty steadily from the higher economic levels down to the lower. The type of parents that produced what were relatively the largest families in 1913 still produced what were relatively the largest families in 1928 ; but the *absolute size of their families* was smaller. Thus, although the correlation between size of family and intelligence remained approximately the same, the standard deviation for size diminished appreciably. Now, if the average number of children in the dullest and most numerous families was smaller in 1928 than in 1913, this would tend to increase the average level of intelligence of the entire sample. Once again, so far as my own meagre evidence goes, I am doubtful whether this factor would produce any *large* increase ; but it would be rash to infer from experience in one area to the probable effects in another.

7.—Another problem that is bound to cause some difficulty is the need to ensure that all children in the area, or in the specified sample, really get tested. *Absenteeism* may be an important factor, because it is the dullest children who are most likely to stay away when a test has to be taken. In London, during the early decades of the century, truancy and irregular attendance were by no means infrequent ; but by the end of the first quarter they had considerably diminished. In the Scottish surveys out of the official number of pupils on the registers in 1932, 9 per cent. did not take the test ; out of those on the registers in 1947, 7 per cent. did not. The drop in the number is too small to be of great importance; but it might be interesting to know whether the proportion of duller children among the absentees had altered.

8.—Several critics have inquired how far *the varying conditions under which the tests were actually administered and marked* can have affected the results. On this point the present volume gives full and reassuring information. Dr. McIntosh, Director of Education for Fife, explains in detail the careful administrative arrangements; Dr. Kennedy-Fraser describes the way in which the individual testing was carried out; and Mr. Maxwell explains how the group tests were corrected and marked. These chapters are excellent, and will be of the greatest service to those who desire to undertake similar investigations. In so large a survey it was, of course, impossible for a single person to be directly responsible for all the testing, and extremely difficult to arrange that the field-workers should be precisely the same in both inquiries. But the fullest precautions appear to have been taken to exclude any serious errors arising from this particular source. At two or three schools, it is said, internal evidence indicated that the tests had not been properly administered; and in other instances quite impossible marks were reported. However, only 240 cases had to be discarded on these grounds—a remarkably small number considering the extent of the inquiry.

9.—Now, so long as we consider each of the foregoing factors by itself, we are likely to conclude that its influence must be far too slight to account for the large difference found. To lay great stress on them would be (as critics quoted in the report have observed), rather like trying to " explain away " an unexpected rise. So far I agree. But I would suggest that the *combined* effect of *all* such factors might easily account for most of the difference, if not for all.

Nevertheless, I doubt whether they would be sufficient to convert a large decline, of the size deducible from the correlation, into an apparent rise, of the size actually observed. I therefore, venture to cite yet another possibility, which I put forward in discussing my own results. " It is conceivable that some kind of reversional or regressional effect may operate, which tends to keep the mean of the population (as well as the variability of the population) rather more constant than the usual methods of computation would imply. Statisticians have repeatedly drawn attention to the fact that certain statistical characteristics of human populations show an unexpected constancy—as if, once a *biological equilibrium* had been reached, the population tends to preserve it, in spite of disturbing changes."[1]

I do not wish to invoke any easy or sweeping postulate about the preservation of " the balance of nature." In particular cases, no doubt, the balance does seem to be kept relatively constant for an appreciable period. Whether in any given population during any given period it has been preserved is a question of fact to be investigated afresh on each occasion. What remains beyond doubt is that no great ingenuity is required to conceive either genetical or ecological processes by which the decline that would ensue from the differential birthrates, when operating in the abstract or by themselves, may be, in any actual and concrete instance, partly or wholly counteracted.

---

[1] I have amplified the suggestion a little in *Intelligence and Fertility* (Hamish Hamilton, 1946), pp. 25-6. Professor Penrose has since suggested a simplified ' model population ' whereby a stable equilibrium could be achieved by assuming (i) a marked degree of assortative mating, and (ii) a much smaller fertility (or even zero fertility) in the homozygotes than in the heterozygotes (*Eugenics Review*, XLI, 1949, pp. 23-4). " The essence of the suggestion," he writes, " is that differential fertility is a phenomenon due to heterosis." Thus, as he adapts it to our present problem, the mechanism described implies that the low fertility of the higher classes is an effect of biological or physiological causes operating quite mechanically to compensate for the decline that would otherwise ensue. That seems difficult to believe. All the evidence indicates that the small size of families produced by the more intelligent is due to voluntary limitation, inspired by economic and personal motives rather than biological.

Most of the tentative suggestions that I have enumerated above have, no doubt, already occurred to the Scottish investigators. Several are discussed or mentioned in their chapters. But in answer to the various questions which the *Report* has provoked, both at educational conferences and in the press, it seemed helpful to catalogue the more likely alternatives that appear to call for further examination. In all probability every one of the factors mentioned may exercise some influence, whether small or large. Hence in future inquiries the statisticians will have the complicated task of attempting to assess the relative importance of each.

So far as the main issue is concerned, the final outcome of the present report is admittedly inconclusive. The one lesson that we can all readily endorse is that drawn by the Royal Commission, after reviewing this additional evidence. " There is an immediate need for further research into the effect of the differential birthrate on the average level of intelligence of the nation ; and we consider that direct investigation in other parts of the country on the lines of the Scottish inquiry, together with research into the respective influences of environment and heredity in intelligence, and the effectiveness of the tests by which these factors are assessed, should be encouraged by the Government."

Meanwhile, the highest praise is due to the Population Investigation Committee for securing funds for the inquiry, to members of the Mental Survey Committee for carrying it out, and above all to its Chairman, Sir Godfrey Thomson, for his part in organizing so elaborate a piece of work. The volume now published avowedly presents only a preliminary account of the main results. More detailed analyses of the vast array of data collected are promised in the near future. But one novel feature in the investigation deserves special commendation. As already noted, the second survey included, besides the actual testing, a ' sociological inquiry.' The schedules of questions used are appended in a suggestive chapter by the late Dr. Miller Young ; and, as Professor Glass points out in the concluding pages of the book, the additional information thus secured should eventually be of the greatest value, not only in throwing light on the main issue, but also in elucidating the general social background against which the work of teachers and educationists has to be carried out.

CYRIL BURT.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES.[1]

A Statistical Study of Teachers' Preferences Among their Pupils.

By D. COCKBURN

(*Durham M.Ed., June*, 1949).

In this study, the preferences of fourteen teachers among twenty-two children were examined. Techniques were used which had been devised by Peel in the study of æsthetic preferences. The teachers' rankings of the children in order of *personal liking* were regarded as the measure of a complex criterion. The children were then ranked in terms of three objective qualities: Intelligence, Smartness, and Ascendancy-Submission. Intelligence was measured by a standardised test, while the other two qualities were measured by averaging the rankings of expert judges, among whom agreement was high. The teachers' preferences were analysed in terms of the objective qualities.

For purposes of comparison, the preferences were examined in four sex-groups, i.e., Men and Women Teachers judging Boy and Girl Pupils. The weightings in various groups are tabulated below.

PREFERENCES	COMPARATIVE WEIGHT OF OBJECTIVE QUALITIES		
Men's for Boys	1. Intellect.	2. Smartness.	3. Ascendancy.
Women's for Boys..	1. Intellect.	2. Smartness.	3. Submission.
Men's for Girls.	1. Smartness.	2. Intellect.	3. Submission.
Women's for Girls..	1. Smartness.	2. Intellect.	3. Submission.

The choice of objective qualities was governed by (*a*) psychological relevance and (*b*) statistical considerations. The data was treated as an analysis of a complex criterion in terms of three determiners, and it was necessary that the determiners should each correlate highly with the criterion, while the correlation between determiners was low.

Four methods were used to analyse the preferences:

(1) Each teacher's preference was expressed as a weighted equation by estimating regressions on the determiners.

(2) A matrix, inclusive, of both preferences and determiners, was analysed and the factor loadings interpreted.

(3) The preference of the teachers as a whole was expressed in terms of the regressions of the factor loadings on the three determiners.

(4) An alternative estimate of the group preference was made, using multivariate analysis.

Within the limitations of the data, it was found that the personal preferences of this group of teachers could be expressed in terms of psychologically meaningful qualities.

---

[1] Theses Outlines must be submitted through the Head of the Department in which the research was carried out.

A Comparative Study of the Methods of Measuring Interest in Science and its Relation to Ability and Achievement.

BY L. ANGUS

(*Abstract of thesis approved for M.Ed., Durham University,* 1949).

Most pupils in the middle forms of grammar schools require to be advised whether their future course should be given a scientific or a non-scientific bias. This cannot always be decided by the results of Term Examinations, for often the most intelligent pupil is equally good at most subjects. Bearing in mind the importance attached to Interest in educational theory, the question arises : " Does a pupil's interest in a subject predict his future success in that subject or not ? "

The aims of this research were : (*a*) To compare methods of measuring Interest in Science ; (*b*) To investigate the relation between Interest, Intelligence, and Achievement with a view to predicting Achievement in Middle School Science by Intelligence plus Scientific Interest, and thus discovering whether tests of scientific interest should be included in any test battery used for educational guidance.

One hundred and twenty boys and girls at a County Durham Grammar School were selected for this investigation. The tests and estimates used were as follows :

(1) Total Science Marks—the average scaled marks of three terminal examinations. (2) New Type General Science Test constructed and item-validated by the writer. (3) Teachers' Science Estimate on a fifteen-point scale. (4) Fife Test No. 2 (Ability for Science). (5) Teachers' Interest Estimate on a five-point scale. (6) Information Test of Interest. (7) Vocabulary Test of Interest. These last two tests were constructed by the writer using the " extended technique of choice " devised by Peel (see this *Journal,* Vol. XVIII, 1948, p. 41). These two tests were item-validated against pupils of the same age at a Northumberland Grammar School. (8) Interest Questionnaire constructed by the writer. (9) Otis Gamma Intelligence Test. (10) Alexander Performance Scale. (11) Peel's Non-verbal Test T.S.8.

The tests were inter-correlated by the Product Moment method when the following correlation coefficients were obtained :

CORRELATION COEFFICIENTS

DECIMAL POINTS HAVE BEEN OMITTED.

	T.Sc.	Sc.E.	N.T.T.	F.T.	Otis.	A.P.S.	T.S.8.	Int.E.	Inf.	Voc.	Qs.
T.Sc.	—	947	772	542	526	485	381	481	343	190	214
Sc.E.	—	—	777	519	553	475	422	421	333	193	156
N.T.T.	—	—	—	530	524	539	438	493	394	190	195
F.T.	—	—	—	—	635	470	406	185	124	-025	243
Otis	—	—	—	—	—	514	488	315	259	029	310
A.P.S.	—	—	—	—	—	—	687	299	185	036	309
T.S.8.	—	—	—	—	—	—	—	189	162	054	227
Int.E.	—	—	—	—	—	—	—	—	803	647	387
Inf.	—	—	—	—	—	—	—	—	—	592	286
Voc.	—	—	—	—	—	—	—	—	—	—	288
Qs.	—	—	—	—	—	—	—	—	—	—	—

T.Sc.	Total Science Marks.	T.S.8.	Peel's Non-verbal.
Sc.E.	Science Estimate.	Int.E.	Interest Estimate.
N.T.T.	New Type Test.	Inf.	Information Test.
F.T.	Fife Test.	Voc.	Vocabulary Test.
A.P.S.	Alexander Performance Scale.	Qs.	Questionnaire.

This matrix was analysed by Thurstone's Centroid Method when the following factor loadings were obtained :

FACTORS EXTRACTED BY CENTROID METHOD OF ANALYSIS

DECIMAL POINTS HAVE BEEN OMITTED.

	Factor Loadings				Communality
	I	II	III	IV	
T.Sc.	823	264	394	-152	925
Sc.E.	811	308	419	-125	944
N.T.T.	795	209	236	102	741
F.T.	602	373	-168	-233	584
Otis	676	277	-232	-150	610
A.P.S.	662	302	-317	297	717
T.S.8.	585	294	-326	266	605
Int.E.	709	-580	125	075	861
Inf.	605	-579	128	131	734
Voc.	393	-620	112	-012	552
Qs.	424	-216	-308	-176	352
Percentage of the total test variance	43·4	15·5	7·41	3·07	69·4

The following factors were postulated :

(1) A General Scientific Ability factor common to all the tests.
(2) A Scientific Interest factor.
(3) A General Intelligence factor.

The fourth factor was very small and was ignored.

Examination of the factor loadings showed that Total Science Marks with a communality of ·925 had the largest factor I loading, and was the best single criterion of scientific achievement. Of the Interest tests, the Vocabulary test was most highly saturated with the interest factor, and was probably the purest measure of interest, but its communality of ·552 indicated large errors and specific factors. The Interest Questionnaire had the least interest factor and lowest communality and was the worst measure of interest. With a communality of ·861 the Teachers' Interest Estimate was seen to be the best criterion of Interest.

It was not found necessary to rotate the factor axes to reveal the distinctive difference between the Achievement and Interest groups of tests.

Validation with Teachers' Interest Estimate as criterion, showed the Information Test to be the most valid and the Questionnaire to be the least valid :

Correlation between Interest Estimate and :

	r	S.E.$_r$
Information Test	·803	·032
Vocabulary Test	·647	·053
Questionnaire	·387	·077

Repeat reliabilities showed the Information Test to be the most reliable and the Vocabulary Test to be the least reliable of the interest tests.

Information Test	$r = \cdot 802$	S.E.$_r$ = ·048
Questionnaire	$r = \cdot 777$	S.E.$_r$ = ·054
Vocabulary Test	$r = \cdot 757$	S.E.$_r$ = ·058

The prediction of a complex achievement criterion by the battery of Interest tests revealed that the Vocabulary Test measured something other than achievement to a greater extent than the other two tests. Similarly, the prediction of a complex criterion of Intelligence by the same battery revealed that the Vocabulary Test measured something other than Intelligence to a greater extent than the other two tests.

The prediction of a complex criterion of scientific achievement by Interest plus Intelligence revealed that Intelligence as measured by the Otis test played a larger part in predicting achievement than did Interest as measured by the Interest Estimate. When, however, tests of Intelligence and Interest were combined, a better prediction of achievement was obtained than by either used separately.

Although the investigation showed that further research is needed to establish a really reliable criterion of scientific interest, it also showed that tests of scientific interest should be included in any battery of tests used for predicting scientific achievement, and thus for educational guidance in the middle forms of grammar schools.

A Consideration of the Ability for Algebra Shown by Secondary School Entrants.

By JAMES T. WILSON

Summary of thesis in part fulfilment of the requirements for the Ed.B. Degree, Glasgow University, 1949.)

All the pupils who passed on from primary education to one senior secondary and one junior secondary school were tested for :

(1) Intelligence, using the Simplex Junior Test (Harrap).
(2) Ability in functional arithmetic, using the calibration test set by the Glasgow Education Department, in May, 1948.
(3) Ability for Algebra, using Fife Test 3 (University of London Press).

The complete sample consisted of 269 pupils—141 boys and 128 girls. The distribution of intelligence in the sample corresponded reasonably well with that expected from the normal frequency distribution. The mean intelligence quotient for the girls (viz., 97·5) was lower than that for the boys (viz., 100·7), but the difference was not significant.

Thorndike's criterion, that an intelligence quotient of 110 is a necessary prerequisite for satisfactory work in algebra, was applied. On this basis it was found that only 30 per cent. of senior secondary pupils could be expected to make satisfactory headway in algebra—and no junior secondary pupils.

Correlations were worked for (*a*) Functional arithmetic and algebra ; (*b*) I.Q. and algebra ; (*c*) combined arithmetic and I.Q. score and algebra.

For this purpose the scores were referred back to the primary classes from which the pupils came. This meant that each group considered had been exposed to the same teaching with regard to arithmetic. The sexes were considered separately and all correlations were worked by the product-moment method. Since the groups were so small, the standard or probable error formulæ were not used to test significance. Instead, the reliabilities of the correlation coefficients were tested against the null hypothesis.

The results obtained were as follows:

BOYS

School	No. of boys.	Correlations		
		Arith. – Alg.	I.Q. – Alg.	Arith.+I.Q.-Alg.
A	15	·836	·899	·922
B	23	·669	·769	·794
C	23	·763	·849	·868
D	22	·814	·707	·802
E	10	·835	·627**	·743*
F	20	·548*	·677	·691
Ga	15	·768	·840	·836
Gb	13	·835	·788	·839

GIRLS

School	No. of girls	Correlations		
		Arith – Alg.	I.Q. – Alg.	Arith.+I.Q. – Alg.
A	12	·790	·692*	·841
B	14	·872	·727	·825
C	18	·902	·895	·929
D	19	·604	·674	·761
E	16	·735	·686	·755
F	23	·878	·769	·856
Ga	12	·716	·680*	·751
Gb	18	·826	·894	·916

In these tables, Ga and Gb denote two classes from the same school. The correlations which are not significant at the 1 per cent. level, are marked with asterisks. Those with one asterisk are significant at the 5 per cent. level and that with the double asterisk fails to attain significance at the 5 per cent. level. It only falls slightly short of this level, however. The required correlation for significance at this level would have been ·632.

Conclusions.—From these results it would seem that there is a marked relation between ability to do functional arithmetic and ability for algebra. There seems to be little to choose between this particular test of functional arithmetic and the intelligence test where prediction of ability for algebra is concerned. It seems, too, that the gain in accuracy obtained by combining arithmetic and intelligence scores is slight.

The boys scored better than did the girls in functional arithmetic and this difference was found to be significant.

According to the Fife Test of Ability for Algebra, only 30 per cent. of the senior secondary pupils would profit by a course in algebra—and no junior secondary pupil. It is interesting to note how closely this agrees with the Thorndike criterion. When we compare the two criteria we find that they are in agreement about 80 per cent. of the pupils, and disagree about the remaining 20 per cent. If we decide to allow pupils who satisfy one or other of the criteria to begin algebra, then 37 per cent. of the boys, and 31 per cent. of the girls in the senior secondary school would be allowed an algebra course, while the subject would not appear in the time table of the junior secondary school. The data give no evidence of any real association between sex and ability for algebra. Boys and girls, therefore, begin the course with equal chances of success. By analysing the answers given by pupils to the Fife Test of Ability for Algebra, the following conclusions were reached:

(1) Almost all pupils passing from the primary to the secondary school can extend simple arithmetical progressions. The senior secondary pupils can handle ascending and descending series with equal facility, but the junior secondary pupils find ascending series easier.

(2) Only the better senior secondary pupils could extend geometrical progressions. Thus, for many pupils, powers are going to prove a source of difficulty. It seems, however, that more use might be made of simple series in teaching new facts in algebra to good pupils.

(3) Many pupils find difficulty in appreciating that $7+5=12$, $12-5=7$, $12-7=5$, and similar sets of results, are merely different aspects of the one relation. This difficulty will occur in algebra when the pupil is required to solve simple equations.

(4) Substitution in formulæ is not a simple and obvious process.

(5) There are signs that the teaching of coding and decoding will prove a helpful introduction to the teaching of substitution in formulæ.

BOOK REVIEWS

Questioning in the Classroom: F. M. AUSTIN. (University of London Press, 1949, pp. 183 and iv. 6s. 6d.)

This book fills a gap in modern literature on teaching methods and should be of considerable use to students training to be teachers. The author, Mrs. Austin, lecturer in Education and Psychology in the University of Birmingham, has had a rich experience in the training of teachers, which she is able to bring to bear on the problems with which she deals. She shows throughout good judgment in evaluating various classroom techniques and sympathetic understanding of the difficulties which confront inexperienced teachers faced with large classes. She writes also with the saving grace of humour.

The first chapter of the book which is concerned with the questions of young children should be of special interest to parents. The following chapters deal effectively and comprehensively with the art of questioning in the classroom. They contain much sound practical advice, on the preparing and the putting of questions, and on dealing with the answers from individuals or the class. Then there follows a more general consideration of the misuses and uses of questioning and of the place and methods of written, as opposed to oral, work.

It may be objected that the topic dealt with is artificially limited in scope. There is throughout a concentration on *verbal* questioning and perhaps, therefore, an undue emphasis on verbalism in thought. It may be argued that the training in the art of thinking which comes through children's first-hand experiences—though their observing, experimenting and attempting to express themselves in other media than words, should not remain unrelated to questioning in the classroom. But, within the limits which she herself sets, the author is able to give much useful advice concerning, and many well selected illustrations of, the technique of questioning in the classroom, together with penetrating analyses of the educational purposes served by the various methods described. O.W.

Testing School Children: WILLIAM STEPHENSON. (London: Longmans, Green and Co., 1949, pp. 127, 12s. 6d.)

Dr. Stephenson's main purpose, so he tells us, is " to give parents and teachers the opportunity of reading about the problems, difficulties, and techniques of selecting children for various schools of the country, whether Public or State." He begins with three short introductory chapters on the practical merits of psychological testing and of ordinary scholarship examinations. He then goes on to discuss in greater detail the extent to which selection should be based on ' personality ' and on ' general ability,' devoting a chapter to each. The rest of his book is chiefly concerned with the use of record cards and of standardized tests of various types—tests of intelligence, of performance, of creative imagination, and of temperament and character.

As a result of his experience both in the Army and in schools, Dr. Stephenson has been led very considerably to modify his earlier views. Formerly he appeared as the most enthusiastic champion of Spearman's two-factor theory. But he now cites evidence, which, he says, seems to " lend support to Professor Thomson's views rather than to Professor Spearman's." He describes himself as " one of those who hold that, even though the child has been formed partly by heredity and partly by environment, heredity is but a spark to the conflagration that constitutes environment . . . As for intelligence, this is an empty concept in itself, which we don't even measure in practice: it should be replaced by Galton's notion of natural ability, with zest and honest endeavour at its roots." Thus, when applying intelligence tests to children of elementary and preparatory schools, he himself has found that the former, at any rate during their earlier years, fall far below the latter. Nevertheless, though he does not cite his reasons, he holds that the elementary children are really " no less innately intelligent," and that their poor performance is due chiefly to " inhibitions."

He has become a great admirer of the Rorschach test. It is, he declares, " a telling probe into the sufficiencies and insufficiencies of personalities " ; it can disclose " the degree to which a child is well-adjusted for its age," and many other suggestive characteristics. But here again the evidence that he offers is far from convincing.

In his final chapter, on " The Principles and Practice of Selection,' he summarizes the earlier papers from the Symposium in this *Journal*. As regards principles, he fully agrees that the official suggestions of the Ministry seem to ignore " the one solid fact known about children at eleven," namely, that, at that early age, their abilities are general rather than specialized. As regards practical procedures, he thinks the plan suggested by Burt is " ideal " and " hard to better " ; but he fancies it would be " difficult to achieve in the hurly-burly of the head master's study ; something more practical is required." He himself concludes that it would be " simplest to elect children into the various secondary schools on the basis of an order of merit list, making it depend a little on a scrutiny of record cards, and having regard to the proportions of *G*, *V*, and *K* entering into the total score."

The whole book is written in a lively and challenging style ; and it would be unfair to criticize the sweeping or dogmatic character of some of its statements as though it was addressed to scientific specialists. One of its most interesting features consists in the plates and appendices of test-materials. These include several ingenious forms of the familiar matrix test, a coloured ink-blot test, which certainly produces far more distinctive and varied responses than a Rorschach blot, coloured ' art-form ' tests, and a good deal of other test-material, designed for children aged eleven or thereabouts, which should certainly prove helpful to those experimenting in psychological methods of selection. C.B.

An Outline of Psycho-analysis: SIGMUND FREUD, translated by JAMES STRACHEY. (The Hogarth Press, 1949, pp. ix+84, 8s. 6d. net.)

This is the last book written by Freud. It contains no fundamentally new doctrine, but aims at a short exposition of the main essentials of the principles of psycho-analysis. It is remarkable for its clarity, and for its conciseness relative to some of Freud's other writings. It is divided into three sections : I—The Mind and its Workings, including a discussion of dream-interpretation and of the development of the sexual function. Part II is devoted to the technique of psycho-analysis. Part III gives " The Theoretical Yield." The book will be of most value not to the beginner, but to the student who has read some of Freud's earlier books and papers, and who wishes to know what he thought most essential and permanent in his doctrine.

Psychiatric Examination of the School Child: MURIEL BARTON HALL, M.D. (Edward Arnold, pp. viii, 368, 15s.).

The author of this book was formerly psychiatrist at the Liverpool Child Guidance Clinic. She explicitly states that her views are based on routine practice ; but she displays a much wider knowledge of the relevant literature than is sometimes found in books on child psychiatry. The many references to the work of others, however, are often merely descriptive, rather than critical.

The subjects covered by the title of the book are treated very thoroughly, and with balanced judgment ; but it should be noted that the title refers only to the *examination* and not the treatment of the children. This is touched on only very slightly, and with surprisingly little reference to the question of discipline.

More Adventures in Reading—Books I and II : GERTRUDE KEIR. (Oxford Univ. Press, 9d. each.)

These attractive readers are specially written for backward readers who have worked through the earlier series of *Adventures in Reading*, using the oral vocabulary and dealing with the common interests of the average ten-year-old child.

Psychological Aspects of Clinical Medicine : STEPHEN BARTON HALL. (H. K. Lewis and Co., p. xi+416, 21s. net.)

After a spate of books by psychiatrists, most of which reveal little knowledge of psychology, or over-emphasize some things to an absurd degree, or consist largely of detailed reports of cases, with unconvincing diagnoses and few references to the outcome of treatment, it is an immense relief to read this volume by the lecturer in psychological medicine at Liverpool University. The book is comprehensive and admirably planned; the discussion is logical and clear, the general attitude cautious and well-balanced. The unusual amount of genuine psychological background makes the book at once a solid introduction for the medical student, and of great value for the non-medical psychologist whose work includes the handling of neurotics.

OTHER PUBLICATIONS RECEIVED

The mention of a book in this list does not preclude a subsequent review.

The Nature and Technique of Understanding : HUGH WOODWORTH (Wrigley Printing Co., Vancouver, pp. 142, 24s.)

The Psychologist in Industry : M. E. STEINER (Blackwell, pp. vii-107, 9s. 6d.)

Developmental Teaching} JAMES L. MURSELL (McGraw-Hill, pp. vii-374, 21s.)

Social Psychology : LAPIERE and FARNSWORTH (McGraw-Hill, xiii-626, 27s.)

The Search for the Beloved : NANDOR FODOR (Hermitage Press, pp. viii-412, 30s.)

Motor Performance and Growth : HAROLD E. JONES (Cambridge University Press, pp. x-181, 24s.).

Looking and Doing : O. GARNETT (Blackwell, pp. xv-143, 4s.)

Modern Concepts of Education and their Significance in the Post-War World : CARLETON WASHBURNE (Education Fellowship—N.E.F. Monograph No. 10, pp. 44, 1s. 6d.)

Police Problems of Today : SIR HAROLD SCOTT (Stevens, pp. 23, 2s. 6d.)

Psycho-Analysis—a Modern Delusion : FRANK KENYON (The Pioneer Press, pp. 141 5s.)

International Congress on Mental Health, London, 1948 (H. K. LEWIS) :

- Vol. I History, Development and Organisation, pp. 154, 10s.
- Vol. II Proceedings of the International Conference on Child Psychiatry (11th—14th August) pp. xiii-142, 10s.
- Vol. III Proceedings of the International Conference on Medical Psychotherapy (11th—14th August), pp. xv-129, 10s.
- Vol. IV Proceedings of the International Conference on Mental Hygiene (16th—21st August), pp. ix-330, 20s.

The Teaching and Learning of Biology in Secondary Schools : T. L. GREEN (Allman and Son, Ltd., pp. 214, 8s.)

The Retarded Child : HERTA LOEWY (Staples Press, Ltd., pp. 52, 3s. 6d.)

The Social and Legal Aspects of Sexual Abnormality : EDWARD GLOVER (I.S.T.D. Publications, pp. 16, 1s.)

Learning Theory in School Situations : SWENSON, ANDERSON and STACEY (Oxford University Press, pp. 103, 9s.)

Interim Report 1948 *of the Children's Reception Centre, Mersham, Kent* (obtainable from the centre, pp. 35, 3d.)

The Golden Years : JOHN BENIENS (Anthroposophical Publishing Co., London, pp. v-185, 6s.)

Body and Mature Behaviour: M. FELDENKRAIS (Routledge and Kegan Paul, pp. viii-167, 12s. 6d.)

Current Trends in Industrial Psychology : a group of eight lectures (University of Pittsburgh Press, pp. 198, 22s. 6d.)

Life in the Nursery School : L. DE LISSA (Longmans, Green, pp. xx-339, 12s. 6d.)

Probation and Re-education : ELIZABETH R. GLOVER (Routledge and Kegan Paul, pp. xii-292).

Psychological Statistics: Quinn McNemar (Chapman and Hall, pp. 364, 27s.)
New Directions in Science Teaching: Anita Laton and Ralph Powers (McGraw-Hill, pp. xi-164, 15s.)
Psychology in Education: Herbert Sorenson (McGraw-Hill, pp. xiv-535, 24s.)
Introduction to Community Recreation: George D. Butler (McGraw-Hill, pp. xiv-568, 27s.)
Psychological Factors in Education: Beaumont and Macomber (McGraw-Hill, ix.-318, 18s.)
First Steps in Number (Books I-IV): Dorothy Williams (Oxford University Press—Geoffrey Cumberlege, pp. 15 each book, 6d. each).
Les Origines du Caractère Chez l'Enfant: Henri Wallon (Pr. Univ. de F., pp. xi-233 360 francs.)
Les Notions Morales Chez l'Enfant: Germaine H. Wallon (Pr. Univ. de F., pp. 250, 500 francs.)
Le Faux Intellectualisme: W. Berteval (Pr. Univ. de F., pp. 120, 200 francs.)
L'Individualisation de l'Enseinement: Henri Bouchet (Pr. Univ. de F., pp. vii-277, 400 francs.)
L'Etre et le Travail: Jules Vuillemin (Pr. Univ. de F., pp. 182, 360 francs.)
L'Univers Morbide de la Faute: Dr. A. Hesnard (Pr. Univ. de F., pp. xiv-469, 700 francs.)
La Musique et l'Homme: André Cuvelier (Pr. Univ. de F., pp. viii-286, 600 fr.)
Philosophie deu Rire et des Pleurs: Alfred Stern (Pr. Univ. de F., pp. 256, 320 francs.)
La Croissance Mentale (L'Enfance): René Hubert (Pr. Univ. de F., pp. xi-357, 600 francs.)
La Croissance Mentale (L'Adolescence): René Hubert (Pr. Univ. de Fr., pp. 359-638, 480 francs.)
Essai sur la Personnalité: Marcel Deschoux (Pr. Univ. de F., pp. vii-385, 500 francs.)
Introduction à la Criminologie: Etienne Greeff (Pr. Univ. de F., pp. 414, 1,000 francs.)
L'Enfance Anormale: M. Prudhommeau (Pr. Univ. de F., vii-156, 200 francs.)
L'Unité de la Psychologie: Daniel Lagache (Pr. Univ. de F., pp. 62, 150 francs.)
L'Influx Nerveux et la Psychologie: Dr. Paul Chauchard (Pr. Univ. de F., pp. 158, 200 francs.)
Pensée Implicité et Perception Visuelle: Jacques Paliard (Pr. Univ. de F., pp. 125, 200 francs.)
Histoire de l'Enseignement en France: Michel Glatigny (Pr. Univ. de F., pp. 127, no price given.)

The Editor regrets that the concluding article in the series, on Educational Research, by Professor Schonell, has not been received in time for this number of the Journal.

VOLUME XX. JUNE 1950 PART 2.

THE IMPROVEMENT OF READING ABILITY.[1]

BY L. B. BIRCH
(*Head of the Remedial Education Centre of the Burton-upon-Trent Education Committee*).

I.—*Introduction.* II.—*A testing programme.* III.—*A remedial centre.* IV.—*Results of a more recent survey.* V.—*Conclusions.* VI.—*Summary.*

I.—INTRODUCTION.

IN the Autumn of 1947 an experiment was started in the County Borough of Burton-upon-Trent to try to improve the standard of reading in the schools. It was feared that, owing to the war, the standard had probably fallen and the experiment was designed to discover the actual state of affairs and to take steps, if necessary, to improve them.

First, a survey of reading ability was carried out with three age groups.[2] Those chosen were the nine, ten and eleven year old groups. They were given a non-verbal group test of intelligence and an individual reading test of the graded vocabulary pronunciation type. Then an estimate of each child's retardation (or otherwise) in reading was obtained, in terms of mental years, by comparing the reading age with the mental age. It was found that approximately 23 per cent. of the tested children had reading ages two years or more below their mental ages. In the secondary modern schools, after the abler third of the population had been creamed off, 40 per cent. were retarded by this amount.

It should be indicated here, perhaps, that none of these children had been evacuated, or suffered except very slightly from air-raids, Burton being a " safe " area during the war.

Since it was estimated that before the war some 15 per cent. of the school population were retarded (Schonell—*Backwardness in the Basic Subjects*, p. 81), a figure above 20 per cent. was cause for alarm. Perhaps it might be pointed out here, however, that other recent surveys, for example that reported by Miss Hammond in August,[3] 1948, show similar conditions.

While the survey work was going on in the schools, heads were asked for their opinions about their pupils' reading skill. Of the ten junior school heads who were consulted, only two admitted that they were at all concerned about any weakness in reading. Two others thought that children coming from the infant departments were not quite as able as they had been ; of the remainder some seemed unaware that there was any reading problem. When the survey was complete it was interesting to note that the most satisfactory position was in the two schools where a weakness was realised. In one of those schools there were only two children retarded and they by only 2 years. In the other only 3 per cent. were retarded more than 2 years.

---

[1] EDITOR'S NOTE :
The Editor has had submitted to him a considerable number of reports on Child Guidance Clinics, which, with regret, he has usually been unable to accept because they have described merely the organisation of the clinic (common to many in this country) and the number of children referred for each of various reasons. Such figures are already fairly familiar from many clinic and other reports. The Editor, however, wishes to make it clear that he welcomes even very brief reports, if they give definite evidence from follow-up studies, or if they deal with some specially selected aspect of the work.

[2] For details, see L. B. BIRCH : " The Remedial Treatment of Reading Disability " : *Educational Review*, Feb., 1949.

[3] " Attainment in Reading " : *Times Educational Supplement*, Aug., 1948.

It was clear from the attitude of the heads concerned that the first step was to hold a general discussion on the implications of the findings of the survey so that all would become aware of the exact nature of the problem and be invited to co-operate. At this stage the Birmingham University Institute of Education was asked to assist by providing a series of lectures on reading disability. During the course an exhibition of reading materials and textbooks was also arranged.

II.—A TESTING PROGRAMME.

Prior to the experiment estimates of a child's reading ability were largely subjective. They were usually obtained by asking him to read aloud from some selected passage and were assessed in "marks out of ten" or in some other arbitrary manner. They were not related to the results of intelligence tests so that, while they may have indicated the child's position relative to his class mates, they did not show what his reading potential might be. The child of very high native ability, for example, who was reading at only an average level, received no special note.

In the experiment it was decided to employ the method used in the survey as a routine procedure to bring into clear focus the child who was failing in reading. Three intelligence tests are now given to all children in Burton during their stay in the primary school. The first of these, The Moray House Picture Test, is given at the age of six-and-a-half. This test is standardised for seven year olds, but it seems adequate for younger children. It gave a mean of 99·6, a standard deviation of 14.5 and was approximately normal in distribution. While no great importance is attached to the exact scores obtained on this test, it is of value in two ways. Firstly, it gives an indication of those children who, by reason of their poor native endowment, are not yet mature enough to be likely to succeed in learning to read, and whose formal reading can with advantage be delayed. Secondly it often shows up the boy or girl of superior ability who, for some reason or other, is not making the progress he should.

At eight years of age the Sleight Non-verbal test is given and its results compared with the findings of the previous testing. Then, at ten plus, a verbal test, the Essential, is given. If, in the second or third testing the result seems at variance with that of a preceding test or with the teacher's estimate, the Terman and Merrill test is given, supplemented, if necessary, by other individual tests.

From the results of these tests a fairly reliable estimate of each child's mental age can be finally made and this can be taken as an approximate guide as to his potential reading ability. The actual reading ability is measured each year, and in some cases each term, with a graded word vocabulary test which measures the ability to recognise and pronounce words divorced from their context. From time to time this is supplemented by tests of reading comprehension. Schonell's tests are used and to a less extent those of Burt, Vernon and Watts. The intelligence tests and the standardised tests of attainment are thus complementary, for if a child has a reading age appreciably below his mental age, he can be regarded as retarded in that subject.

III.—A REMEDIAL CENTRE.

Many children who are shown by the tests to be retarded, especially those who are fairly young and those whose disability is slight, can be helped towards a complete recovery in their own schools by a little extra attention from their class teacher. There are others, however, whose disability is so severe that they need special remedial teaching, either individually or in small "opportunity groups." To help these children an educational remedial centre was set up in the Borough in 1947. Here small groups from all the schools, graded as to their actual and potential reading ages, are given special help to overcome their difficulties and catch up with their more fortunate class mates. Each child attends the centre for about an hour daily, the rest of his time being spent in his own school. At the week-end the periods are changed round so that groups come at different times in consecutive weeks.

The methods used in the centre are largely those of the enlightened classroom. After the initial aversion which such children almost always feel towards a " backward " subject has been broken down by free activity and play methods, the job of learning to read is started all over again. There is plenty of stimulating material, carefully graded so that each child can experience the quick success necessary to sustain him in his task. Before attendance at the centre, a careful case history is taken so that special needs can be catered for. Close contact is also kept with each child's home and school.

When reading age approximates to mental age, attendance is discontinued. This may take anything from three to eighteen months, or more. Almost all cases respond to treatment, though sometimes the old apathy creeps back again when the novelty of the centre has worn off. Such cases are usually those whose unfortunate home background or severe emotional upset call for more extensive treatment than we are at present able to give.

IV.—A MORE RECENT SURVEY.

In October, 1949, just two years after the first survey, a similar investigation was carried out with one age group to see whether there was any improvement. The children tested were those just entering the secondary modern school at the age of eleven plus. The following table shows a comparison between the two surveys.

COMPARISON BETWEEN RESULTS OF READING RECOGNITION TESTS GIVEN TO COMPARABLE GROUPS OF SECONDARY MODERN CHILDREN.

Year	Number of Children	Average Age	Average Reading Age	Percentage Retarded 2 Years or more
1947	337	11·5 yrs.	8·9 yrs. ($\sigma=2·5$)	Approx. 40 per cent.
1949	332	11·5 yrs.	10·1 yrs. ($\sigma=1·4$)	Approx. 16 per cent.

The two groups being considered here are, with negligible exceptions, those who failed to obtain a place in a selective secondary school. As can be seen from the next table, they represent similar proportions of their respective age groups and are not significantly different in general ability. It was, unfortunately, not possible to give the same intelligence test to each group, but each worked a Moray House Intelligence Test as part of the Special Place Examination at the age of eleven, and though these were different tests, it is known that they correlate highly with each other.

Year	Number in Whole Age Group	Percentage of Age Group going to Sec. Mod. Schools	Mean I.Q. of Whole Age Group on different Moray House Tests	Standard Deviation on Moray House Tests
1947	614	66	101·5	15
1949	601	64	100	14

The Graded Word Vocabulary Test standardised by Schonell[1] which was used in both surveys, was unknown to almost all the teachers in the town in 1947, but since then it has been used frequently. There was thus opportunity for coaching to have taken place. (The effect of practice without coaching on this type of test has been shown to be very small.)[2] In order to see whether there had been any coaching one child in five, taken at random, was given Vernon's Test[3] which it is almost certain was unknown in the schools. There was an average difference of less than ·2 of a year between the two testings, so that it may be assumed that the higher score of the 1949 group was not due to familiarity with the test.

The figures giving the percentage of retarded children were obtained by taking the difference between reading age and mental age. Mental ages were calculated in 1947 from one non-verbal test—the Jenkins, and in 1949 from the higher of two test scores on Sleight and Essential Tests. They can thus only be regarded as approximate, though it is unlikely that the true position would be less favourable than is shown as the procedure gave a slightly higher I.Q. in 1949 than 1947.[4]

V.—Conclusions.

It would seem that there has been a considerable improvement in reading ability in the two years under review. The main causes of improvement seem to have been:

(1) Stimulation of interest in the junior schools on the problem of reading retardation.
(2) An adequate scheme of testing designed to throw into sharp focus the child who is failing.
(3) Remedial work being carried out in schools and in the remedial centre.
(4) Generally improved conditions in schools and homes due to settling down after the war together with availability of better textbooks and other materials.

That the improvement in reading ability is largely due to our interference and not to a general improvement in conditions is suggested by the fact that twenty-four Secondary Modern school children who were found to be very retarded in 1947, but who could not be given any treatment till 1948, were found to have made almost no progress during that year.

VI.—Summary.

1.—A survey of reading ability showed some 20 per cent. of the children in schools in Burton-upon-Trent had reading ages two or more years below their mental ages. This was a worse state than seems to have existed before the war.

2.—The following steps were taken to improve matters:

(*a*) Teachers were made aware of the magnitude of the problem and their interest stimulated by lectures, discussions, exhibitions, etc.
(*b*) A scheme of testing was introduced which would make the detection of the failing reader more certain.
(*c*) A remedial education centre was set up to deal with the worst cases.

3.—A recent survey of children entering the secondary modern school at eleven plus, showed that their average reading age had risen from 8·9 years in 1947 to 10·1 years in 1949 and that the number of children retarded two years or more had fallen from approximately 40 per cent. to approximately 16 per cent.

---

[1] F. J. Schonell: *The Psychology and Teaching of Reading* (Oliver and Boyd).

[2] "Remedial Treatment of Reading Disability": *Educational Review* (Vol. 1, No. 2).

[3] P. E. Vernon: *The Standardisation of a Graded Word Reading Test* (Univ. Lond. Press).

[4] By this method the mean I.Q. in 1947 was 89·5 and in 1949 it was 91·5.

THE ESTIMATION OF DIFFICULTY OF VOCABULARY.

By P. E. VERNON
(Institute of Education, University of London)

I.—*Introduction.* II.—*Construction of a quality scale.* III.—*The value of the quality scale.* IV.—*Summary of conclusions.*

I.—INTRODUCTION.

EDUCATIONAL psychologists often criticise the difficulty of vocabulary in school textbooks,[1] or in Army instruction[2], in Government propaganda, and other types of teaching. They point out that students, whether children or adults, cannot be expected to follow historical, scientific, technical or other matter if they do not know the meanings of a lot of the words. The same problem arises with educational broadcasts to schools or to the Forces, which are often written by experts who have little acquaintance with the limited capacities of their audiences. Thus Wall[3] reports that a number of words used in recent Forces broadcasts could not be defined by considerable proportions of listeners. He concludes that vocabulary, which is literary (e.g., 'succinctly, stilted'), or which embodies abstract conceptions (e.g., 'relevant, theorising') is likely to give difficulty, and that the best yardstick lies not in the relative length of words, but in concreteness and familiarity in colloquial use.

It may well be that this feature of verbal matter has been over-emphasised at the expense of difficulties in sentence construction and style, abstractness of subject matter and so forth, as Flesch[4] and Gray[1] suggest. But it presumably has some importance in view of the extremely poor vocabulary ability of below average and dull students. How then is the difficulty or familiarity of words to be assessed? Evidence is given below that the judgments, even of intelligent teachers, are somewhat inconsistent and inaccurate. In an investigation by Kirkpatrick and Cureton[5], the combined ratings by five English instructors of 100 words correlated ·71 with the difficulty of these words when embodied in a multiple-choice vocabulary test. This sounds promising, but the average correlation for any one judge would be only ·51 to ·59.[6]

American educationists and psychologists generally resort to the Thorndike-Lorge word list, which gives the frequency of occurrence of some 30,000 words derived from counts of many millions of words in English classics, current literature, newspapers and other sources. Davis[7] disagrees with this as an index of familiarity, and quotes correlations as low as ·10 between Thorndike frequency and the empirically determined difficulty of vocabulary items. As Kirkpatrick and Cureton point out, however, Davis's items were designed to test precision rather than range of vocabulary ; that is, their difficulty resided more in the multiple-choice responses provided than in the words themselves. These authors found correlations of ·47 to ·56 between the difficulty of more straightforward items and frequency. Even these figures indicate that the word list, like the average judge, can provide only a very rough estimate of ease or difficulty.

---

[1] Cf. FLEMING, C. M.: *Research and the Basic Curriculum* (University of London Press, 1946). GRAY, W. S., *et al.*: *Reading in General Education.* (Washington, D.C.: American Council on Education, 1940).

[2] Cf. VALENTINE, C. W.: *Principles of Army Instruction* (Aldershot: Gale and Polden, 1942).

[3] WALL, W. D.: "Broadcasting for the Backward" (*Times Educ. Suppl.*, 1949, Sept. 22).

[4] FLESCH, R.: *The Art of Plain Talk* (New York and London: Harper, 1946).

[5] KIRKPATRICK, J. J. and CURETON, E. E.: "Vocabulary Item Difficulty and Word Frequency" (*J. App. Psych.*, 1949, 33, 347-351).

[6] This estimate assumes that the mean correlation between the judges lies between ·4 and ·6.

[7] DAVIS, F. B.: "The Interpretation of Frequency Ratings Obtained from 'The Teachers Word Book'" (*J. Educ. Psych.*, 1944, 35, 169-174).

Naturally, the Thorndike frequencies are still less likely to be appropriate in Britain. An English edition of the Thorndike Junior Dictionary has been published by Ballard and Palmer,[1] which omits many peculiarly American words and substitutes some which are more commonly used in this country. But no fresh frequency counts of British usage have been carried out.

There appears to have been one attempt only to supply a vocabulary based on actual ease of definition. Dale and Chall[2] list the 3,000 commonest words known by at least 80 per cent. of American Fourth Grade pupils. Its American origin renders it unsuitable for use in this country, and familiarity among 10 year olds can hardly be considered equivalent to familiarity among adults. Actually, only one of the hundred words tried out below by the writer occurs in this list (namely ' old-fashioned '), although twenty-four of the hundred were known by 75 per cent. or more of his adult groups.

It occurred to the writer that teachers, psychologists and others might be able to make more accurate estimates with the aid of a quality scale, consisting of a series of specimen words of known difficulty. The following experiments were, therefore, undertaken as a preliminary step in an investigation of the factors making for intelligibility of educational broadcasts.

II.—CONSTRUCTION OF A QUALITY SCALE.

One hundred words which seemed likely to cover the range from moderate to great (though not very great) difficulty were chosen. Most of these were taken from the scripts of recent Forces Broadcasts. Five lists of twenty each were duplicated with the following instructions :

Write down in your own words the meaning of each of the following :

Example : House—a building that people live in.

The lists were well shuffled and were given to 380 regular recruits to the Royal Navy. Thus an average of seventy-six men answered each list. Twenty minutes were allowed, this being sufficient for almost all the men to attempt all the words they thought they could do.

It may be objected that ability to write a definition would not provide a good criterion of understanding, since many of the duller recruits would have difficulties in expressing themselves on paper. However the scoring was very lenient. Full credit was given to any explanation, however ungrammatical or badly spelled, which suggested a grasp of the essential meaning. Partially right definitions, or explanations which merely illustrated the word in a particular context, were given Half Credits. For example :

UNIT	*Full Credit* :	One, or one section. A self-contained item.	*Half* :	A group of Service men. A meaurment. Part of electricity.

---

[1] BALLARD, P. B., and PALMER, H. E. : *The Thorndike Junior Dictionary* (London : University of London Press, 1947).

[2] DALE, E., and CHALL, J. S. : " A Formula for Predicting Readability" (*Educ. Res. Bull.*, Ohio State University, 1948, 27, 11-20, 37-54).

PETITION	Were people put their names down for something they want. Paper for a reason with peoples names on it.	A dockument or paper of peace. A poll organised without authority. To sign your name for a reprieve.

Table I lists the percentage frequencies of correct (plus ½ of the half-correct) responses. It should be remembered that the groups were not large, hence these figures are somewhat unstable. In other similar groups, some of the words might be 10 per cent. or so higher or lower.[1]

TABLE I.

PERCENTAGE PASS-RATES AMONG MALE ADULTS FOR 100 WORDS.

Per cent.	
1	Arbitrary, Periphery.
2	Apathy, Philanthropist.
4	Depredation.
5	Exploitation, Perfunctory.
6	Remuneration.
9	Debility.
10	Impartial.
12	Relevant, Vocation.
13	Attributed to, Sporadic.
14	Sequel.
18	Proximity.
25	Substance.
26	Principles.
34	Omen, Quartz.
35	Contend with, Mesh, Universally.
37	Excess, Theory.
38	Unit.
41	Optimistic.
43	Devaluation.
45	Contrasting.
47	Flattering, Negotiation, Premature.
49	Presumably, Research, Unique.
50	Canvassing, Fruitless, Productivity.
51	Alloy, Petition.
52	Reliability.
53	Analysis, Automatically, Measurement.
54	Prediction.
55	Behaviour.
56	Vitamins.
57	Finance, Vigorous.
58	Efficient, Inevitable.
59	Compressed, Evidence, Parallel.
60	Traditional.
61	Illusion.
62	Capacity, Emphasise.
63	Fertile, Source, Visibility.
65	Concerned with.
68	Expansion, Manufacturer.
69	Composed of, Secondary importance, Superior.
70	Equivalent, Horsepower, Survival.
71	Approximately, Battery, Experiment.
72	Anæsthetic.
73	Treacherous.
74	Accelerate.
75	Prohibit.
76	Refugee.
77	Tractor.
78	Imagine.
80	Fulfil.
81	Concentrate on, Majority.
83	Inflammable, Old-fashioned.
84	Skilful.
85	Abolish.
86	Complicated.
87	Generous.
88	Agriculture.
90	Uncertainty.
92	Bankrupt, Temporary.
93	Expensive, Take Heed.
94	Expert.
95	Commence, Employer, Immediately.
97	Similar.

The recruits had taken the standard Admiralty intelligence tests, hence it was possible to state that they were very close to the average for the general male 18 year population. They were, however, considerably restricted in range at both ends. On the assumption (which was supported by earlier item analysis investigations) that the tetrachoric correlation between ability to pass the average word and verbal intelligence is close to ·4, it was possible to predict the item pass-rates in a group representative in scatter as well as in mean intelligence. Pass-rates between 80 and 20 per cent. were found to be almost unaffected, and words of 90 or 10 per cent. difficulty were altered only by 1·3 per cent. But very difficult words, with pass-rates of 1 per cent. among our recruits, would be passed by approximately 5 per cent. of a group of normal scatter ; and very easy words with 99 per cent. pass-rates would be failed by some 5 per cent. of such a group. Thus the percentages quoted in Table I can be taken as reasonably representative for 18 year male adults, except at the extremes.

---

[1] The Standard Errors are ±2·5 per cent. for words of 95 or 5 per cent. difficulty, ±3·5 per cent. for words of 90 or 10 per cent., ±5·0 per cent. for words of 75 or 25 per cent., and ±5·8 per cent. for words of 50 per cent. difficulty.

Lest it be thought that these figures are on the high side, perhaps because the recruits might be of superior ability, the following check was made. The numbers of words in the Thorndike dictionary at various frequency levels are known, and from the results of our recruits on our 100 words we can determine their pass-rates on specimen words at each Thorndike level. Thus for words of 2-3 Thorndike frequency, the mean pass-rate was 31 per cent, 4-6 level 51 per cent., and so on. We can, therefore, estimate the total words known by the average recruit, though only roughly, since the words selected at the extremes were far from typical of the most and least difficulty in the dictionary. Nevertheless, it is safe to say that the estimate is close to 9,000 out of the 18,000 commonest words in the English language, and according to Terman[1], this is precisely the average vocabulary at 14 years Mental Age. True the Average Adult vocabulary is supposed to be 11,700 words, but it is generally recognised that the old Stanford-Binet test overestimated the average adult intelligence level[2]. Thus this calculation indicates that the percentages in Table I are certainly not over-optimistic.

Yet another check is provided by the unaided judgments of forty-four teachers, who assessed 16, 27, 33 and 24 per cent. of the words in our list as falling within the four Grades described below. According to our results, the proportions should have been 15, 23, 38 and 24 per cent. respectively. Although some judged the words as more difficult, some less so, than our findings, yet the overall proportions agree quite closely.

The words were now divided into four grades, and specimens were taken to provide the following quality scale of difficulty.

Grade 3.—Words likely to be known only by about the top 10 per cent. of the adult population (say those with School Certificate or better) :
Arbitrary, Attributed to, Debility, Depredation, Exploitation, Impartial, Periphery, Philanthropist, Proximity, Remuneration, Sporadic, Vocation.

Grade 2.—Likely to be known only by one quarter to one-third of the population :
Devaluation, Excess, Flattering, Fruitless, Negotiation, Omen, Premature, Presumably, Principles, Substance, Theory, Unique.

Grade 1.—Likely to be known by half or more, i.e., by the average adult :
Anæsthetic, Analysis, Automatically, Behaviour, Concerned with, Equivalent, Expansion, Inevitable, Manufacturer, Survival, Traditional, Vitamins.

Grade 0.—Likely to be known by three-quarters or more :
Agriculture, Bankrupt, Complicated, Expert, Expensive, Fulfil, Imagine, Majority, Old-fashioned, Prohibit, Skilful, Tractor.

The correspondence between our percentages and Thorndike frequencies are of considerable interest. Tetrachoric correlations were obtained at the 84th,

---

[1] TERMAN, L. M. : *The Measurement of Intelligence* (London : Harrap, 1919).

[2] Cf. VERNON, P. E. : " A Study of the Norms and the Validity of Certain Mental Tests at a Child Guidance Clinic " (*Brit. J. Educ. Psych.*, 1937, 7, 73-82).

63rd and 24th percentiles, i.e., at the dividing lines between the above grades. The coefficients were ·78, ·60 and ·30. This means that there is quite close agreement between frequency and difficulty for words of above average difficulty, but very little among more common and easier words. The words which were more difficult than would be expected from their frequency of written usage include:

> Attributed to, Theory, Unit, Principles, Universally, Substance.

As Wall suggests, they tend to be abstract or literary words which are difficult for the average and dull adult to define. Words which are more often known than would be expected from their frequency include:

> Productivity, Devaluation, Canvassing, Visibility, Anæsthetic, Vitamins, Bankrupt, Refugee.

These seem to be *either* concrete terms from engineering, medicine or commerce, *or* terms which frequently appear in the press or in broadcasts.

III.—THE VALUE OF THE QUALITY SCALE.

Three lists of thirty-three words were selected from the 100, parallel as to difficulty. Twenty-two grammar school teachers or training college lecturers engaged on a Higher Degree course assessed one or other list without the quality scale. Fourteen others assessed a list with the aid of part of the scale, i.e., with sets of eight words from each Grade taken from the lists which they were *not* being asked to assess. On comparing the two sets of assessments with the ' true ' Grades, the results of the two groups were very similar. Tetrachoric correlations between assessments and ' true ' grades ranged from ·28 to ·96 for different teachers, and the medians were ·68 and ·71 in the two groups. Clearly the results were too much affected by the ability of the particular teachers who happened to be in the two groups to be reliable.

In a further experiment, therefore, a group of twenty-two graduate students, who were training as teachers, each rated one list without, and a different list with, the aid of the quality scale.

A definite improvement in standards of grading was obtained. Since each list contains 5 Grade 3, 7 Grade 2, 13 Grade 1, and 8 Grade 0 words, the *total* grades should always add up to 42. At the first assessment 7 students only gave approximately correct total grades, ranging between 35 and 49; whereas 15 students' total grades were more than 7 points too high or more than 7 points too low. At the second assessment, when guided by the quality scale, 15 gave total grades correct within ±7, and 7 were still seriously inaccurate.[1]

Taking all 80 sets of gradings, the deviations from correct total grades are shown below (deviations above and below 42 were roughly equally divided):

	No. of Judges	Deviations from Correct Total.				
		±3	±4 – 5	±6 – 10	±11 – 15	±16 and over
		per cent.	per cent.	per cent.	per cent.	per cent.
Without Scale	44	16	14	32	27	11
With Scale . .	36	36	22	22	11	8

---

[1] This improvement is moderately significant statistically. Chi-square = 5·81 (or 4·46 with Yates's correction); P = < ·05 > ·01. A *t* test gives a similar probability.

It will be seen that the standards of judges working without the scale are often considerably too high or too low, and though a few of those using the scale are also inaccurate, the difference is significant at P=·01.

Correlations among the students were closely similar to those among trained teachers. When all 726 assessments were plotted in a single scatter, product-moment r (without Shepherd's correction) was ·567. A similar scatter for the 726 assessments given with the aid of the scale yielded r=·609. None of the conventional tests of significance will apply when both persons *and* judgments are involved; but it seems unlikely that this improvement is any greater than would be produced merely by the more accurate total grades. In other words, the quality scale does help slightly in pinning the judgments to correct standards, but not in grading relative order of difficulty. Given a passage of prose, the judge who uses the scale should be able to assess the total grades of its more difficult words somewhat better, though he may make just as many errors in deciding which particular words are very or only moderately difficult.

The percentage of correct gradings for the average word rose by only a small amount, from 44·9 per cent. without the scale to 47·8 per cent. with it. The number of words out of 99 which were graded correctly by at least 6 out of 7 students rose from 8 to 14, and the number misgraded by as many students dropped from 20 to 15. But these differences are not significant.

Thus, the usefulness of the quality scale is somewhat disappointing. Nevertheless, the present writer is finding it a valuable standby in assessing the vocabulary level of a large number of educational broadcasts. It is possible also that persons who are not teachers, and have less experience of immature vocabularies, would be more greatly assisted by it. In judging possible Grade 3 words, the scale might be supplemented by consulting the Thorndike Junior Dictionary. Almost all words of 1, ½ or zero frequency, except those likely to be more familiar in Britain than in America, can fairly safely be called 3. But the dictionary cannot be regarded as a reliable guide to Grade 2, 1 or 0 words.

IV.—SUMMARY OF CONCLUSIONS.

1.—Vocabulary level is often regarded as an important factor in difficulty of educational materials, but cannot at present be satisfactorily assessed. Neither Thorndike word-book frequencies, nor estimates of difficulty by single teachers, correlate more than ·5 to ·6 with ease of definition of words.

2.—Definitions of one hundred specimen words were obtained from groups of 76 male recruits, whose average intelligence was close to that of the total population. On this basis, the words were divided into four Grades, ranging from Grade 3—known only by top 10 per cent.—to Grade 0—known by at least 75 per cent. Sample words from each Grade are presented as a Quality Scale, with the aid of which it is hoped that better estimates of the difficulty of other similar words may be made.

3.—In an experiment where student teachers were asked to assess the difficulty of some of the words, without or with the quality scale, there was scarcely any improvement in the accuracy of their relative gradings. But the absolute standards of assessment showed a significantly closer approximation to the correct grades.

AN INVESTIGATION OF THE VALUE OF HIGHER SCHOOL CERTIFICATE RESULTS IN PREDICTING PERFORMANCE IN FIRST-YEAR UNIVERSITY EXAMINATIONS.

BY EIRIAN M. WILLIAMS
(*Assistant Lecturer in Education, Sheffield University*).

I.—*Purpose of the enquiry.* II.—*Scope of the enquiry.* III.—*Procedure.* IV.—*Results.* V.—*Analysis of results.* VI.—*Discussion of results.* VII.—*Conclusion.* VIII.—*Summary.*

I.—PURPOSE OF THE ENQUIRY.

EACH year the number of candidates applying for admission to the universities increases more rapidly than the number of places available. The problem of selection thus becomes continually more acute. One basis of selection is the candidate's performance in the Higher School Certificate examinations. This account describes an attempt to examine the extent to which Higher School Certificate results in any subject are predictive of the first-year university examination results in that or any other subject.

II.—SCOPE OF THE ENQUIRY.

The investigation was concerned with students sitting first-year examinations at Sheffield University in the Faculties of Arts, Science and Medicine during the years 1943-47 inclusive, who had taken the Higher School Certificate examination of the Northern Universities' Joint Matriculation Board in the July preceding their entrance to the university. Candidates of other examining boards were omitted in order to minimise the error from possible variations in standard in the Higher School Certificate examination. Students who for any reason did not proceed immediately to the university after sitting the Higher School Certificate examination, were excluded from the enquiry in order to avoid the complication of a gap in their academic studies. Altogether 938 students sat first-year examinations in the three Faculties concerned during the years 1943-47, inclusive, but of these only 495 came within the scope of the enquiry.

III.—PROCEDURE.

The various possibilities of pairing Higher School Certificate subjects with first-year university subjects were listed and the students taking each combination counted. Thus the student who took Higher School Certificate Principal Chemistry, Principal Physics and Principal Pure Mathematics and proceeded to sit Intermediate Chemistry, Physics and Mathematics, would make the following combinations: Principal Chemistry with Intermediate Chemistry, Principal Chemistry with Intermediate Physics, Principal Chemistry with Intermediate Mathematics, Principal Physics with Intermediate Chemistry, and so on. Other students provided examples of different combinations according to Faculty. Points which arose at this stage were dealt with as follow:

(i) In the Faculty of Arts some subjects can be taken at the ordinary Intermediate standard or at Advanced Intermediate standard (for those proceeding to Honours). These were dealt with separately.

(ii) In the Faculty of Science students exempted from the Intermediate examination take the First Year Qualifying Examination. These two examinations were treated separately. For this purpose it was decided to include First M.B. Chemistry and Physics with the corresponding examinations of the Faculty of Science as being of the same standard. With minor differences the papers are the same in both Faculties.

On this basis the number of actual combinations totalled over 700. However, even after adding together the number of cases over the period of five years, only 68 of the subject combinations provided anything like large enough numbers for statistical examination. Correlation coefficients were calculated in the 68 cases between Higher School Certificate examination marks and university examination marks. The error introduced by considering the five years as a unit will be discussed later.

IV.—Results.

Tables 1, 2 and 3 show:

(i) The coefficients of correlation obtained (r).

(ii) The probable error of these coefficients (P.E.$_r$).

(iii) The number of students involved in each calculation (n).

Correlation coefficients in square brackets are less than three times their probable error.

V.—Analysis of results.

Preliminary examination of these results brings out two points:

(i) The values of the correlation coefficients are, on the whole, low; 30 out of 68 are less than three times their probable error. Correlation coefficients for Subsidiary Higher School Certificate marks follow the same general trend as those for Principal Higher School Certificate marks.

(ii) The correlation coefficients between Science subjects are generally higher than those between Arts subjects.

Tables 1, 2 and 3 also show some rather unexpected correspondences. For example, Principal Geography marks correlate more closely with those in Intermediate Modern History than with Intermediate Geography marks. Such differences between correlations have, however, been considered in relation to their probable errors.

For reasons already noted, figures for a five-year period were combined. This is obviously a source of inaccuracy. However, in some of the few cases where the numbers were large enough to permit the calculation of correlation coefficients for separate years, this was done. In general, the correlation coefficients for the individual years were within the limits indicated by the probable errors of the correlation coefficients in Tables 1, 2 and 3. However, either in 1942 or in 1943, in whichever year the intake of students into any particular university department was well below the average in quality, the correlation coefficient was well below the average. These low correlations were the exception and the single correlation for the five-year period appeared to be sufficiently reliable to provide a basis for an introductory investigation.

VI.—Discussion of results.

(*a*) *General discussion.* An attempt was made to identify some of the factors underlying the correlations. The relevant Higher School Certificate and University examination papers were studied and were also discussed along with marking schemes and courses with members of the university staff and teachers preparing pupils for the Higher School Certificate examinations.

TABLE 1

CORRELATION COEFFICIENTS FOR PRINCIPAL H.S.C. SCIENCE SUBJECTS

First-Year University Examination Subjects	H.S.C. Examination Subjects														
	Principal Biology			Principal Chemistry			Principal App. Maths.			Principal Pure Maths.			Principal Physics		
	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.
1st M.B. Biology	·77	·06	20	·56	·08	38	—	—	—	—	—	—	·47	·09	34
First-Year Qualifying Chemistry	—	—	—	·50	·06	74	—	—	—	·31	·09	48	·47	·06	68
Intermediate Chemistry	[·24]	·10	37	·44	·06	97	—	—	—	[·13]	·10	41	[·20]	·07	89
First-Year Qualifying Maths....	—	—	—	·47	·06	71	·56	·08	32	·61	·05	75	·34	·06	88
Intermediate Maths.	—	—	—	[·11]	·15	41	—	—	—	·46	·08	35	·29	·09	47
First-Year Qualifying Physics	—	—	—	·50	·05	93	[·33]	·13	23	·42	·06	80	·44	·05	105
Intermediate Physics	[–·09]	·13	25	·33	·07	87	—	—	—	·38	·09	43	·33	·07	87

TABLE 2

CORRELATION COEFFICIENTS FOR PRINCIPAL H.S.C. ARTS SUBJECTS

First-Year University Examination Subjects	H.S.C. Examination Subjects														
	Principal English			Principal French			Principal Geography			Principal History			Principal Latin		
	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.
Intermediate Economics	[·09]	·09	40	[·00]	·12	20	[·24]	·10	37	[·22]	·10	39	—	—	—
Intermediate English	·33	·05	124	[·07]	·07	97	[–·01]	·12	31	·29	·07	69	[·15]	·13	25
Advanced Intermediate French	[·16]	·12	30	·39	·09	48	—	—	—	—	—	—	—	—	—
Intermediate French	·44	·05	104	·68	·04	76	[·13]	·08	49	[·06]	·08	78	·58	·13	14
Intermediate German	[–·04]	·16	19	[·13]	·17	16	—	—	—	—	—	—	—	—	—
Intermediate Geography	—	—	—	—	—	—	·48	·07	64	[·17]	·12	30	—	—	—
Intermediate Modern History..	·24	·07	85	[·05]	·09	56	·52	·07	46	·40	·06	95	—	—	—
Intermediate Latin	[·16]	·07	79	·52	·05	85	—	—	—	[·16]	·09	50	·79	·06	21

TABLE 3

CORRELATION COEFFICIENTS FOR SUBSIDIARY H.S.C. SUBJECTS.

First-Year University Examination Subjects	H.S.C. Examination Subjects																	
	Subsidiary Chemistry			Subsidiary Pure Maths.			Subsidiary English			Subsidiary French			Subsidiary History			Subsidiary Latin		
	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.	r.	P.E.$_r$	n.
Intermediate Chemistry ...	[·35]	·13	20	[·37]	·13	20	[–·14]	·13	25	—	—	—	—	—	—	—	—	—
Intermediate Physics	·56	·11	27	—	—	—	–·66	·08	24	—	—	—	—	—	—	—	—	—
Intermediate Economics ...	—	—	—	—	—	—	—	—	—	·37	·12	23	—	—	—	—	—	—
Intermediate English	—	—	—	—	—	—	—	—	—	[·20]	·13	23	[·15]	·13	27	·36	·08	49
Intermediate French	—	—	—	—	—	—	—	—	—	·70	·05	45	[–·01]	·16	17	·69	·06	33
Intermediate Mod. History	—	—	—	—	—	—	—	—	—	[·03]	·12	32	—	—	—	[·29]	·11	33
Intermediate Latin	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	·76	·04	49

It is obvious that there are many factors which must be considered; but some of these cannot be associated more with one subject than with another, and others do not affect the general picture. In the first place the examining panels of the Joint Matriculation Board undergo periodic changes so that variations in marking might be expected. However, in view of the rigorous standardisation by the Board of all marking, the possibility of variations in Higher School Certificate marking may be regarded as negligible.

Changes in the composition of school staff from year to year and differences in teaching as between school and school introduce another possible factor. Intensive coaching by one teacher may enable less gifted pupils to reach a standard in the Higher School Certificate examinations which they are not able to maintain without this help at the university. Similarly, teaching which is not scholarly may produce poor Higher School Certificate results from potentially good candidates who do better under university conditions. Where it operates this factor will tend to lower correlation coefficients, but without further evidence beyond the scope of this enquiry there is no reason to suppose that its effect is greater on one subject than on another.

A further general factor is the sum of the effects of the transition from school to university life. Generally speaking there is a change from the small Sixth Form class with its opportunities for individual attention to the large university class and university methods which place more responsibility for his own progress on the student than is the case in most schools. Some students may have already assumed this responsibility at school and others entering the university may have developed to the stage where they are ready to do so. These students are more likely to maintain their standard of achievement than those on whom university work makes demands with which they are not sufficiently mature to cope. The new discipline imposed by university conditions is probably a large factor in reducing the general level of correlation coefficients but does not explain the differences between them.

There is one other suggested general factor which should be mentioned. The university lecturer usually knows members of his classes personally and is familiar with the standard of their work throughout the year. He is thus in a position to adjust examination marks of students on the borderline accordingly. The Higher School Certificate examiner does not know the candidates, but to a debatable extent the teachers' estimates which are consulted by the Joint Matriculation Board, enable the general level of a candidate's school work to be taken into account in the same way.

(*b*) *Discussion with reference to specific subjects.* In the attempt to identify the specific factors which may be operating in the case of any particular subject it seemed most profitable to consider the university subjects separately and to take them in the same sequence in which they were listed in Tables 1 and 2.

(i) *University Science subjects.* For *First M.B. Biology* the predictive values of the Higher School Certificate subjects are in the following order: Biology, Chemistry, Physics. Examination of the First M.B. Biology papers served to underline the results. These papers bear a striking resemblance to those set for Principal Higher School Certificate Biology, and their emphasis is on Morphology and Anatomy so that the knowledge of Chemistry needed is slight and of Physics almost negligible. Furthermore, the mathematical element which predominates in the Principal Physics papers is entirely absent from those of First M.B. Biology. In view of this is seems reasonable to suggest that the difference in subject-matter covered by the three Higher School Certificate subjects is the main factor determining the comparatively high correlation coefficient for Principal Biology and the relatively low values for Principal Chemistry and Principal Physics.

The subject matter of the *First-Year Qualifying Chemistry* course is in essence the same as that of Principal Chemistry except that in the former there is increased concern with Physical Chemistry. This side of the subject demands greater knowledge of mathematical techniques and permits freer use of symbolic representation. In spite of this, Principal Chemistry correlates better with First-Year Qualifying Chemistry than does Principal Pure Mathematics, though not significantly better than does Principal Physics. It would appear, however, that the predictive value of Principal Chemistry for First-Year Qualifying Chemistry is largely the result of the basic similarity in the subject-matter covered by the two examinations.

The correlation coefficient for Principal Chemistry is, however, not high and several explanations can be offered for this. In the first place the great emphasis on the Physical Chemistry in the First-Year Qualifying examination would seem to give an advantage at this stage to those who had previously done most Physics or Mathematics. Support to this contention is given in the case of the Mathematics by the figures in Table 4.

TABLE 4

(*a*) H.S.C. subject taken	Average First-Year Qualifying Chemistry Marks		
	All students having taken (*a*)	Students having taken (*a*) and Principal Pure Mathematics.	Students having taken (*a*) and *not* Principal Pure Mathematics.
Principal Chemistry (74 students)	57·3	59·5 (47 students)	53·6 (27 students)
Principal Physics (68 students)	56·9	58·4 (44 students)	54·0 (24 students)
Principal Pure Maths. (48 students)	59·1	59·1	—

In the case of the corresponding figures for Physics (Table 5), the numbers of students not having taken this subject at Principal Higher School Certificate standard before proceeding to First-Year Qualifying Chemistry are, unfortunately, very small.

TABLE 5

H.S.C. subject taken (*a*)	Average First-Year Qualifying Chemistry Marks		
	All students having taken (*a*)	Students having taken (*a*) and Principal Physics	Students having taken (*a*) and *not* Principal Physics
Principal Chemistry (74 students)	57·3	57·1 (67 students)	59·4 (7 students)
Principal Physics (68 students)	56·9	56·9	—
Principal Pure Maths. (48 students)	59·1	58·4 (45 students)	70·0 (3 students)

As far as they go the marks in Table 5 suggest that to have taken Principal Physics gives a student no advantage in the First-Year Qualifying examination in Chemistry. The Principal Physics papers, however, contain a large proportion of numerical problems, and it is, therefore, likely that the Physics marks are some indication of the candidates' mathematical ability. This may help to explain why Principal Physics correlates with First-Year Qualifying Chemistry almost as well as does Principal Chemistry.

Another factor tending to lower the correlation coefficient between Principal Chemistry and First-Year Qualifying Chemistry is suggested by a consideration of the marking schemes for the two examinations. In Principal Chemistry the practical examination carries 12 per cent. of the total marks, whereas at the First-Year Qualifying stage it carries 33⅓ per cent., including marks for a compulsory question on practical work in one of the theory papers. Furthermore, the marking of the First-Year Qualifying Chemistry papers is stiffer in at least one important respect. A question is considered as a whole and not in sections, so that the mere statement of principles, without the evidence that they are understood and can be correctly applied, earns no marks. In Principal Chemistry, by contrast, marks are allotted separately to each portion of a question. It is often the case in the First-Year Qualifying Chemistry papers that the scope of a question is very wide and the candidate must show judgment in confining himself to one aspect of a problem and in selecting his material accordingly. The more closely defined limits of a Principal Chemistry question make this less necessary and may not allow the better candidate the opportunity to reveal himself.

With regard to *Intermediate Chemistry*, there are four main facts requiring explanation :

(i) The very low correlation coefficient for Principal Biology.
(ii) The small negative correlation coefficient for Subsidiary English Literature.
(iii) The low correlation coefficients for Principal Chemistry, Pure Mathematics and, in particular, for Physics with Intermediate Chemistry as compared to the corresponding ones with First-Year Qualifying Chemistry.
(iv) The size of the correlation coefficients for Subsidiary Chemistry and Subsidiary Pure Mathematics.

Between the Principal Biology and Intermediate Chemistry papers there is only a little overlap of subject-matter, and the choice allowed in the Biology papers enables the candidate to avoid the physiological questions where the overlap occurs. Furthermore, the position as regards Sixth Form Biology probably varies more from school to school than does that of any other subject.

Many boys take Biology for the first time in the Sixth Form and, though they may have a good foundation of Chemistry and Physics for the Physiological topics, they often have a good deal of ground to make up in Morphology and Anatomy. Because of the time factor their Principal Biology marks may not represent their potentialities, which may reveal themselves at the end of another year. Many girls, on the other hand, take Principal Biology in the Sixth Form, but do no Chemistry until they come to the university. Their Intermediate Chemistry marks may then tend to be low compared with those of poorer biologists who have previously done Chemistry. In fact, the number of Intermediate Chemistry students who have not done Principal Chemistry is small, among those considered, but as far as they go the figures in Table 6 support the above contention :

TABLE 6

H.S.C. subject taken with Principal Biology	Average Principal Biology Mark	Average Intermediate Chemistry Mark
Principal Chemistry .... (32 students)	53·6	50·5
Subsidiary Chemistry ...... (5 students)	66·8	45·0
No H.S.C. Chemistry .... (2 students)	66·5	39·0

The negative correlation of Subsidiary English Literature with Intermediate Chemistry is in keeping with the generally accepted idea that the Sciences and the "English" subjects require in the main different abilities, and involve very different interests.

In considering the low correlations of Principal Physics and Pure Mathematics with Intermediate Chemistry, the question of mathematical knowledge recurs. The Intermediate Chemistry papers are very similar to those for Principal Chemistry and questions requiring mathematical treatment are few and can be avoided. The mathematician would appear to have no advantage in Intermediate Chemistry, as is suggested by the following figures:

TABLE 7

H.S.C. subject taken (*a*)	Average Intermediate Chemistry Marks		
	All students having taken (*a*)	Students having taken (*a*) and Principal Pure Mathematics	Students having taken (*a*) and *not* Principal Pure Mathematics
Principal Chemistry (97 students)	49·8	49·9 (31 students)	49·7 (66 students)
Principal Physics (39 students)	48·1	46·0 (36 students)	49·5 (53 students)
Principal Pure Maths. (41 students)	45·3	45·3	—

Thus the non-mathematical character of the Intermediate Chemistry papers seems to furnish some explanation of the low correlations of Intermediate Chemistry marks with those of both Principal Physics and Pure Mathematics.

To explain the low correlations of Principal Chemistry with Intermediate Chemistry in view of the similarity of the papers, it is necessary to look for other factors. The proportion of marks allotted to practical work at Intermediate standard is again greater than at H.S.C. standard, being 33⅓ per cent., as against 12 per cent., but this is true also of the First Year Qualifying examination which correlates better with Principal Chemistry. One other suggestion seems to deserve consideration. Many students doing Intermediate Chemistry are those who have not done so well at Higher School Certificate standard (the better ones being allowed to follow the course for the First Year Qualifying examination). Much of the Intermediate course appears to be repetition of H.S.C. work and some students over-estimate their ability, in particular they show poor judgment in using their knowledge and do badly in the examinations.

The correlations with Subsidiary subjects are also interesting. Subsidiary Chemistry provides the same degree of prediction as Principal Chemistry though the marks obtained in Intermediate Chemistry are lower. This is indicated by the following figures:

TABLE 8

H.S.C. subject taken	Correlation Coefficient with Inter. Chemistry	Average Inter. Chemistry Mark
Principal H.S.C. Chemistry	·44	49·6
Subsidiary H.S.C. Chemistry	·35	39·1

Subsidiary Chemistry papers are very similar in approach to those of Principal Chemistry. Two factors, however, are probably responsible for the lower correlation coefficient of Subsidiary Chemistry with Intermediate Chemistry. In the first place there is no practical test at the Subsidiary stage and secondly the choice of questions allowed make it possible to avoid any on Organic Chemistry. This is not possible in Principal Chemistry or at Intermediate standard.

Although the correlation coefficient for Subsidiary Pure Mathematics with Intermediate Chemistry is not significantly greater than zero (mainly the result of a large probable error owing to small numbers), there is a fair chance that it is significantly better than the correlation coefficient of Principal Pure Mathematics with Intermediate Chemistry. It is likely that those students who sat Subsidiary Pure Mathematics intended to specialise in Chemistry or some science subject other than Mathematics. In general, the Subsidiary Pure Mathematics group would contain the better chemists. This is to some extent supported by the figures in Table 9.

TABLE 9

H.S.C. subject taken (*a*)	Percentage of students having taken (*a*) and Principal Chem.	Average H.S.C. Chemistry Mark	Average Inter. Chemistry Mark
Principal Pure Maths.	76	51·0	45·2
Subsidiary Pure Maths.	85	52·5	47·1

It may be that the Subsidiary Pure Mathematics students proceeding to Intermediate Chemistry take the subject more seriously and, being already ahead in the work, present in their examination results a truer estimate of their scientific ability than do the Principal Pure Mathematics group. If the Subsidiary Pure Mathematics papers measure some ability needed for Intermediate Chemistry the correlation for this subject would be raised. It seems impossible at this stage to offer any more definite suggestion than this.

The amount of correlation between *First-Year Qualifying Mathematics* marks and those of the Higher School Certificate subjects considered, apart from Physics, is substantial. The correlation coefficient for Applied Mathematics is slightly lower than that for Principal Pure Mathematics and this is possibly owing to the fact that Applied Mathematics questions can be neglected in the First-Year Qualifying examination. Pure Mathematics questions cannot be avoided, and so the papers give an advantage to students who are better at Pure Mathematics and penalise those who are better at Applied Mathematics.

It has proved impossible to suggest a satisfactory explanation for the higher correlation of Principal Chemistry with First Year Qualifying Mathematics by comparison with that of Principal Physics, especially as the Principal Chemistry papers contain only a small percentage of numerical questions while 75 per cent. of the Physics questions require mathematical treatment.

In order to account for the lower correlations of Principal Pure Mathematics with *Intermediate Mathematics* as compared with First Year Qualifying Mathematics one can only repeat the suggestions made for the same situation in Chemistry. The Intermediate Mathematics students are the weaker mathematicians of the year's entry, and their performance is likely to be upset by new conditions of work, though the degree of disturbance will vary according to the ability and previous training of the student.

It is difficult to point confidently to a reason for the very low correlations of Principal Chemistry and Principal Physics with Intermediate Mathematics. Both are lower than the corresponding correlations with First Year Qualifying Mathematics. However, whatever the extent to which Principal Chemistry or Principal

Physics may test ability to succeed in Intermediate Mathematics, performance in this subject is bound to be influenced by the amount of Mathematics previously done. Analysis of the groups of students proceeding from Principal Chemistry and Principal Physics to Intermediate Mathematics show them to be very heterogeneous. In both cases they comprise students who took the following :

(i) Principal Pure Mathematics and Principal Applied Mathematics.
(ii) Principal Pure Mathematics only.
(iii) Principal Pure-and-Applied Mathematics.
(iv) Principal Pure Mathematics and Subsidiary Applied Mathematics.
(v) Subsidiary Pure Mathematics and Subsidiary Applied Mathematics.
(vi) Subsidiary Pure Mathematics only.
(vii) Subsidiary Applied Mathematics only.
(viii) No H.S.C. Mathematics.

Groups (vii) and (viii) are not found among students proceeding to First Year Qualifying Mathematics from Principal Chemistry or Principal Physics. It seems doubtful, however, whether the difference in composition of the populations is sufficient to account for the differences in correlation coefficients.

The values of the correlation coefficients for *First-Year Qualifying Physics* indicate three points requiring discussion :

(i) The correlation with Principal Physics is not very high.
(ii) The correlation with Principal Applied Mathematics is lower than that with Principal Pure Mathematics.
(iii) The correlation with Principal Chemistry is at least as high as that with Principal Physics.

Where the sections of the subject examined are concerned, the papers for Principal Physics and First-Year Qualifying Physics present a certain similarity. It appears to be a practice of many schools to omit much of the electrostatics and sound. It is possible to avoid questions on these topics at Principal standard and again in the First-Year Qualifying examination. This is not possible in the December terminal examination in the university, in which no choice of questions is allowed and which should, therefore, correlate less closely with Principal Physics. In fact, the correlation coefficient for Principal Physics and the December terminal is ·40 (P.E. = ·06) which is not significantly lower than that for First Year Qualifying Physics. Owing to the nature of the practical work in the university, it is not related to the theory and this may be a disturbing factor to some students and will tend to lower the correlation of Principal Physics with First-Year Qualifying Physics. Furthermore, at the University level it is difficult to set practical tests and marks are awarded on the year's practical work. They form a larger proportion of the total marks than that allotted to the practical paper in Principal Physics. The type of question set at the two levels and the method of marking also contribute to a lower correlation. At the First-Year Qualifying level questions require more discussion, there are fewer numerical riders and, as in Chemistry, questions are marked as a whole and not sectionally.

Principal Applied Mathematics is seen to correlate rather less well with First-Year Qualifying Physics than does Principal Pure Mathematics. It may be that First-Year Qualifying Physics requires the techniques of Pure Mathematics rather than the knowledge of Applied Mathematical principles.

It is interesting that the best prediction for First Year Qualifying Physics is provided by Principal Chemistry. In the Principal Chemistry papers there are few numerical riders and the wording of the questions is more comparable with that of First-Year Qualifying Physics than is that of Principal Physics. It may be that the usefulness of the Principal Physics papers could be enhanced by improvement in these respects.

Much of what has been said to explain the rather low correlation of Principal Physics with First-Year Qualifying Physics also applies in the case of *Intermediate Physics.* Avoidance of electrostatics is, however, not possible in the

Intermediate Physics examination and in view of foregoing statements, this may be a factor in the lowering of the correlation between Principal Physics and Intermediate Physics. The general remarks made with reference to Intermediate students in Chemistry and Mathematics apply here also and, in particular, a more critical type of question and the need to bring theory and practical work into relationship for themselves will have a greater effect at the Intermediate level on the poorer students.

These same factors may operate in lowering Principal Chemistry and Principal Pure Mathematics correlations for Intermediate Physics as compared with those for First Year Qualifying Physics.

It seems reasonable to suggest that the lack of correlation between Intermediate Physics and Principal Biology is almost entirely due to the completely different subject matter of the two examinations and to the character of the questions. The Physics questions are mathematical and involve discussion and criticism in contrast to the description required by the Biology papers.

Very little can be said about the correlation of Subsidiary Chemistry with Intermediate Physics which is the highest for this university subject. Unfortunately, the number of students concerned was so small that the value of the correlation coefficient is doubtful. However, investigation of the previous history of these students shows the group to be most heterogeneous. Between them at the H.S.C. examination they sat for almost every possible Arts and Science subject at all standards, from Subsidiary First-Aid to Principal Latin, and one is prompted to remark that an examination in any scientific subject would give a good indication of their ability to do Intermediate Physics.

The high negative correlation of Subsidiary English Literature with Intermediate Physics suggests that here we have two examinations measuring two entirely different types of ability or of interests apparently rarely occurring in any marked degree in the same person Up to a point this may be the case, but the figures may also underline the tendency of many boys' schools in particular (especially for the sake of scholarships) to encourage a one-sided development of their pupils while paying lip-service to the ideal of a general education.

A general picture which develops from this discussion appears to indicate that Higher School Certificate marks predict performance in the First-Year Qualifying examinations better than that in Intermediate examinations. In all cases the correlation with the First-Year Qualifying examination is higher than the corresponding correlation with the Intermediate marks. The differences are not completely significant but the general trend suggests that the Higher School Certificate examination in Science is fulfilling its function of selecting students for their capacity to undertake advanced work.

Another striking feature of the results for university science subjects is the comparative reliability of the Principal Chemistry marks as an indication of the potentialities of any student. It may be that this provides a clue to the directions in which papers generally might be improved.

(ii) *University Arts subjects.* For *Intermediate Economics* all the correlations are very low except that with Subsidiary French. The predictive value of the latter is almost certainly better than that of Principal English or Principal French and probably better than that of Principal Geography or Principal History. It is difficult to comment usefully on these correlations in the absence of one with Higher School Certificate Economics for comparison. The figures obtained suggest that tests of the abilities required for the other Arts subjects are useless as guides to performance in Intermediate Economics. It was not possible to correlate performance in Intermediate Economics with that in any science subject. The correlation for Subsidiary French remains a mystery and seems of no help in the main problem as all the students involved took Arts subjects.

All the correlations for *Intermediate English* are low, including that for Principal English. This seems an unsatisfactory state of affairs, but it is perhaps unavoidable owing to the difficulty of marking English objectively. It seems generally agreed among those teaching the subject in schools and at the university that good teaching coupled with a fairly high level of general intelligence and conscientious work enable a candidate to do well in Principal English. It is contended that the nature of the subject and the difficulties of marking make it impossible to distinguish on this examination any but the very exceptional. The Principal English examination does not indicate the presence of the critical judgment required for Intermediate English. It may well be that the kind of approach made in the schools to set books is an important factor in the problem. To a certain extent this idea is supported by a comparison of the correlations with Intermediate English of Principal and Subsidiary Latin. That for Principal Latin is not significantly different from zero while Subsidiary Latin gives a better correlation than Principal English. The course for Subsidiary Latin requires no set books while in Principal Latin questions on set books carry slightly less than half the marks.

With regard to the low correlation coefficient for Principal English and *Advanced Intermediate French,* one can only refer again to the remarks made in connection with Principal English and Intermediate English. If those contentions are true they would apply here again since a critical approach to French Literature is particularly called for at this stage. In addition, there is the French language paper which has no counterpart in Principal English.

There is then the problem raised by the lower correlation of Principal French with Advanced Intermediate French than with ordinary Intermediate French. This difference appears to be due to a number of factors which may be broadly summarised as differences in method of approach and varying emphasis on the parts of the subject. The Principal French examination comprises three papers. Paper I, which is definitely linguistic, carries 60 marks; Paper II, which is a literature paper, carries 28 marks; and the oral test, again linguistic, carries another 12. The weighting of these marks is clearly in favour of linguistic as opposed to literary studies. The ordinary Intermediate examination in French is a replica of the H.S.C. examination with the difference that language and literature are given equal weight in marks. However, any decisions concerning students on the borderline are governed by their performance in the language side. Furthermore, the set books for the ordinary Intermediate examination cover exactly the same periods as those studied for Principal French. The case of those taking Advanced Intermediate French presents a different picture. The course involves six hours per week in lectures as compared with three hours for the ordinary Intermediate students. Much of the work the students do is not examined at the Advanced Intermediate stage. In this category is Old French which is also a new subject to them. The problem of planning their private study so as to fit in work on this non-examined material may prove disturbing to students. To some, also, the Classical period in literature studied at the university may be new and in any case the approach is more philosophical and literary than anything to which they have been accustomed. Literature questions set at Advanced Intermediate standard also require a very detailed knowledge of the text, whereas the choice of questions allowed and their framing in Paper II of the Principal French examination enables a candidate to avoid this detailed knowledge. Taken together all these factors appear to provide a likely explanation of the differences in the correlation coefficients.

For *Ordinary Intermediate French* the predictive values of the Higher School Certificate subjects are in the following order: Subsidiary French, Subsidiary Latin, Principal French, Principal Latin, Principal English, Principal Geography, Principal History and Subsidiary History. The consideration given above to the similarity between Principal French and Intermediate French seems to point to a reason for the comparatively high correlation coefficient obtained. The probable importance of the linguistic element in these two examinations seems supported by the high correlation with Intermediate French of Subsidiary French, which provides the same degree of prediction as Principal French though the candidates do less well, and for which no literature paper is set.

The paper for Subsidiary Latin is purely linguistic, and this may be the reason for the high correlation of this examination with ordinary Intermediate French. It may be also that the detailed knowledge of a text which is required for Principal Latin and does not form an important element in Intermediate French is the factor which lowers the correlation between these two subjects by comparison.

It is not clear why Principal English should correlate better with Intermediate French than with Intermediate English. The difference is almost completely significant. However, owing to the emphasis on the linguistic side of Intermediate French, conscientious learning probably takes a student further in this subject than in Intermediate English. It is likely, for this reason and in view of the nature of the examination in Principal English as previously discussed, that marks in this subject are a better indication of performance in Intermediate French than in Intermediate English.

Neither Principal Geography nor Principal History marks shown any correlation with those of Intermediate French. There are possibly two main reasons for this, briefly, the nature of the subject matter involved and the greater difficulty of marking, particularly in History.

For *Intermediate German,* both Principal English and Principal French give correlations not significantly greater than zero. Unfortunately, small numbers prevented the consideration of H.S.C. German. The Principal English papers are exclusively concerned with literature, whereas in Intermediate German there is one Literature and one Language paper carrying equal marks, and the literature paper includes translation, though from prescribed texts. Furthermore, decisions concerning students on the borderline are based on their performance in one section of the language paper which involves translation from English into German. These differences in approach to the two subjects may account for the small negative correlation between Principal English and Intermediate German. The low correlation of Principal French with Intermediate German lends no support to the idea of general linguistic ability.

For *Intermediate Geography* the predictive value of Principal Geography is not very high but is almost certainly better than that of Principal History. The reason seems clear. The Principal Geography course comprises mainly a series of social and regional studies whereas in Intermediate Geography it is the scientific basis of the subject which is being dealt with. The Principal History course thus has no subject matter in common with Intermediate Geography, and the approach is entirely different.

In view of the scientific approach required for Intermediate Geography, it is likely that those students who had taken Principal Geography with science subjects for Higher School Certificate would do better in Intermediate, though they might not have done so at Principal standard. In this way the correlation between the two examinations would be lowered. The figures in Table 10 lend some support to this view.

TABLE 10

Group of subjects in which Principal Geography was taken by candidates	Average Principal Geography mark	Average Inter. Geography mark
Arts	59·5	46·0
Science	53·6	49·8

For *Intermediate Modern History*, the predictive values of the Higher School Certificate subjects are in the following order: Principal Geography, Principal History, Subsidiary Latin, Principal English, Principal French and Subsidiary French.

The low correlations of Intermediate History with English, French and Latin are probably in the main the result of the vastly different content of the courses, but in the case of the English the difficulty of marking this subject objectively will be a major factor affecting the correlation.

There appear to be fairly clear reasons for the rather low correlation between Principal History and Intermediate Modern History again concerned with differences in subject matter. The Intermediate Modern History course deals only with European history whereas in Principal History one paper is concerned with English history. Students proceeding to Intermediate may have gained their highest marks on this English history paper and may, therefore, not do as well at the Intermediate stage. Furthermore, the most popular period of European history for Higher School Certificate seems to be that from 1648—1914, whereas the period studied for Intermediate Modern history is that from 1494—1815. Candidates who happen to have studied the latter period for Principal History will have an advantage at the Intermediate stage. Both these factors are likely to affect the weaker candidates more and will thus operate to lower the correlation between Principal History and Intermediate Modern History.

As the course for Intermediate Modern History is concerned with European History a knowledge of Geography is important. The social and regional character of the Principal Geography course taken in conjunction with the points noted above concerning Principal History may sufficiently explain why Principal Geography correlates better than Principal History with Intermediate Modern History.

The correlation of Principal Latin with *Intermediate Latin* is the highest of those obtained. This is almost certainly due to the close similarity of the papers in both examinations and the nature of the questions which makes marking reliable. As already mentioned in another connection, the questions on set books in Latin demand a detailed knowledge of the texts which it is possible to assess objectively. Although the paper for Subsidiary Latin contains no questions on set books, the marks correlate well with those for Intermediate Latin. This fact could be adduced to confirm that the questions on set books require in general the same kind of approach as the purely linguistic questions and that their absence does not alter radically the order of the candidates.

The lower correlation of Principal French with Intermediate Latin may appear for two reasons. In the first place the literature questions in Principal French, apart from some which can be avoided, do not require the detailed knowledge of the text needed for those on Latin set books, and moreover it is probably more difficult to mark these French literature questions objectively.

The reasons for the low correlations of Principal English and Principal History with Intermediate Latin seem fairly clear. The three courses cover entirely different subject-matter and the approach is different in each case. Again both History and, more particularly, English are difficult to mark objectively.

A survey of the Arts subjects confronts one with a wider variety of subject-matter and of approach than does a similar survey of Science subjects, and many of the Arts subjects present great difficulty in the marking. It is probably for these reasons that the general level of correlation coefficients for Arts subjects is lower than it is for Science subjects.

VII.—CONCLUSION.

In the first place it seems worth while to stress what may appear obvious. The 495 candidates covered by the enquiry made an astonishing number of subject combinations. This indicates the varied composition of the Higher School Certificates obtained as a prelude to any particular university course. Many facts emerging from the enquiry suggest that the subjects taken at Higher School Certificate are not always the most suitable from the point of view of the student's future studies.

In attempting to come to any conclusions from this investigation it is important to bear in mind the functions of the Higher School Certificate and the Intermediate examinations, respectively. Generally speaking the aim in a Higher School Certificate examination in any subject is to arrange the candidates in a strict order of merit in that subject. In any first-year university examination in Sheffield the aim is rather to place the student into one of three groups without concern as to his position within the group. In this way there are separated out for each subject those who have passed well enough to proceed to Special Honours, those who have passed well enough to proceed to a General Degree, and those who have failed. In view of their differing functions, which affect the marking, one cannot expect a strikingly high level of correlation between Higher School Certificate and the first-year university examinations.

It is not claimed that final conclusions can be drawn from this work and several lines are evident along which the investigation should proceed. Larger numbers of students entering the University in the last two years and a greater uniformity in the subject combinations taken for Higher School Certificate should make it possible to obtain correlation coefficients for two years separately, not only with Principal and Subsidiary subjects, but also with those taken at Scholarship level. If these were to indicate the same trends, they might still be peculiar to Sheffield and should be checked in other universities which receive a large proportion of Joint Matriculation Board candidates. This should show whether the correlation coefficients represent differences inherent in the subjects rather than in the courses at a particular university and the methods of marking adopted there. It is hoped to continue the enquiry along these lines.

At the present time much consideration is being given to the framing of Advanced papers for the General Certificate of Education. These, in conjunction with the Scholarship papers, will provide a basis of selection for university entrance. It would appear that the Higher School Certificate papers in some subjects, as constituted and marked at present, do not provide an adequate prediction of the candidate's ability to do advanced work, at any rate in Sheffield University. A thorough investigation of the whole matter might indicate improvements which could be incorporated in the papers of the new examination.

VIII.—SUMMARY.

1.—The enquiry concerned students at Sheffield University who sat first-year examinations in the Faculties of Arts, Science and Medicine in the years 1943-47 inclusive, and who had taken the Higher School Certificate examination of the N.U.J.M.B. in the July immediately preceding their entrance to the university.

2.—Correlation coefficients were calculated between the candidates' marks in each Higher School Certificate subject and their marks in any subject which they studied in their first year at the university. Owing to the great diversity of subjects taken for Higher School Certificate, it was possible to calculate only

68 correlation coefficients since some combinations between Higher School Certificate and university subjects were made by very few students.

3.—Two outstanding facts emerge:

(i) 30 out of the 68 correlation coefficients are less than three times their probable error.

(ii) Correlations between Science subjects are higher than those between Arts subjects.

High positive correlations include values of ·79 for Principal Higher School Certificate Latin and Intermediate Latin, ·77 for Principal Biology and First M.B. Biology, ·58 for Principal Latin and Intermediate French and ·56 for Principal Chemistry and First M.B. Biology.

Low values include ·33 for Principal English and Intermediate English, ·33 for Principal Physics and Intermediate Physics, ·07 for Principal French and Intermediate English, ·13 for Principal French and Intermediate German, and ·13 for Principal Pure Mathematics and Intermediate Chemistry.

Interesting *negative values* are —·14 for Subsidiary English and Intermediate Chemistry and – ·66 for Subsidiary English and Intermediate Physics.

4.—Possible variations in Higher School Certificate marking, change in school staff, differences in teaching and the transition from school to university life are discussed as factors affecting all subjects equally.

5.—Factors affecting specific subjects are discussed in detail with reference to each university subject.

6.—The respective functions of the Higher School Certificate and the first-year university examinations are discussed. Lines are indicated along which it is hoped to continue the enquiry and these may reveal facts which could have some bearing on the framing of papers for the General Certificate of Education.

DIFFERENCES IN INTELLIGENCE AND CULTURAL-EDUCATIONAL LEVEL BETWEEN DIFFERENT EDUCATIONAL GROUPS

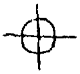
By E. ANSTEY, R. F. DOWSE, and M. DUGUID
(Three members of a Psychological Research Unit working in London).

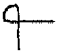
I.—*Description of tests used.* II.—*First experiment.* III.—*Second experiment.* IV.—*Conclusions.*

I.—DESCRIPTION OF TESTS USED.

THE results in this paper are the by-product of an experiment held for a different purpose, namely obtaining norms for a general classification test, hereafter called the G.C. Test. This test has been developed for the purpose of collecting information about the relative intelligence and cultural level of different grades of Civil Servants, particularly the clerical, sub-clerical and manipulative grades, which contain the bulk of the staff. The test will never be used as a selection test ; instead, it will be given to representative groups for research purposes only, with a view to assessing the standard of the group, not of any individual in it. The uses to which this information will be put have not yet been finally settled. Similar general classification tests, however, have been used by the fighting services for some years for such purposes as confirmation of job analysis impressions, and formulation of an overall selection plan for the total employment field, based on estimates of the supply of human material and the requirements of different grades. In the Civil Service some such yardstick would be particularly useful, since the same selection test is seldom used more than once.

The G.C. Test consists of four parts, with separate time limits, the whole test taking precisely one hour to administer.

Part 1 is a pictorial " Code Completion " test, making use of the same fundamental idea as the National Institute of Industrial Psychology's test 70/1, but of entirely original design. Five key pieces are given ; for example, number 1 is like this : 

The test consists of 144 incomplete drawings such as this : 

The subject has to write under each drawing the number of the key piece to which it belongs. This test was constructed by Miss S. C. Topsfield, now lecturer in psychology at Exeter University College.

Part 2 follows the same plan as the very successful Army instructions test, but contains entirely new material in the form of personal particulars. The test involves checking, classifying, coding and filing of these particulars.

Part 3, a mixed verbal intelligence test, contains four types of item :

(i) Abstractions items such as ' stop go up down in * * * '
(ii) Implications items such as ' Which two of the following five words are necessarily implied by the key-word Bicycle ?
(1) Wheels. (2) Bell. (3) Pump. (4) Pedals. (5) Road.'
(iii) Contradictions. Five statements are given, and the subject has to say which two are incompatible.
(iv) Mixed problems.

The time limit for this part is generous, but the later items are distinctly hard.

Part 4, the general information test, was described in the November, 1948, number of this *Journal*. It contains a list of the names of 120 famous people, each of whom has to be matched against the profession or occupation in which he or she became principally famous. The test is intended to be a measure of general cultural-educational level, for which reason sportsmen and film stars were deliberately excluded, but most other professions and branches of knowledge are covered.

II.—FIRST EXPERIMENT.

With the kind co-operation of the War Office, the test was first tried out on a group of 500 soldiers in July, 1947, and scores in the four parts were correlated. The inter-correlations were as follows (correct to two places of decimals):

	Part 2	Part 3	Part 4
Part 1	·63	·61	·52
Part 2	—	·72	·72
Part 3	—	—	·71

The Army " General Service " battery of five tests, and three other tests were given to the same group. The full table of inter-correlations, and the results of a simple summation factor analysis are given in the Appendix. The two main practical conclusions from this experiment were:

(*a*) The fact that the general factor loading for each Part is at least ·76, and the communality at least ·67, suggests that each Part of the G.C. Test is reasonably reliable, despite the short time limits.

(*b*) Judging by their second factor loadings, Parts 2 and 3 are more akin to Part 4 than to Part I. The same conclusion is suggested by the reduced table of correlations above, since Parts 2, 3 and 4 have higher correlations with each other than they have with Part I. Part 4, although ostensibly a test of acquired knowledge, might be expected to have quite a high general factor loading, since more intelligent people tend to acquire more information than less intelligent people.

These correlations, and the factor contents of the four parts, would no doubt be very different for a better educated group.

III.—SECOND EXPERIMENT.

A second experiment with the general classification test was carried out in September, 1948, primarily for the purpose of obtaining norms. For these 500 soldiers, mostly recruits of 18-19, the Personnel Selection Officers' ratings for ' Educational Standard ' were available. Educational standard is a rating on an eight-point scale. 1 means roughly university education; 2, higher school certificate or equivalent qualification; 3, school certificate; 4, full-time education up to 16, but without school certificate; 5, full-time education up to 15 or normal education but very good scores in the Army educational tests; 6 means normal education up to 14; 7, educationally retarded; and 8, illiterate. Combining the smaller groups, out of 500 soldiers, 32 were ES 1 or 2, 68 ES 3, 54 ES 4 or 5, 237 ES 6, and 109 ES 7 or 8. The following tables show the distributions of scores in separate parts of the test obtained by these five educational groups. To facilitate comparison, the figures have been converted into percentages, each column showing the percentage of members of that ES obtaining various scores. The 90th, 50th and 10th percentile scores are also given at the foot of each column. Percentages and percentile scores have been rounded off to the nearest whole number.

TABLE 1

DISTRIBUTION OF SCORES IN PART 1.

	ES 7 and 8	6	4 and 5	3	ES 1 and 2
No. of Cases	109	237	54	68	32
Range of Scores	% of cases	% of cases	% of cases	% of cases	% of cases
135—144	—	—	2	6	16
120—134	—	—	6	9	25
105—119	—	2	7	12	9
90—104	1	5	15	25	28
75—89	2	16	29	24	13
60—74	9	19	15	14	9
45—59	22	24	7	6	—
30—44	32	21	13	3	—
15—29	19	12	6	1	—
0—14	15	1	—	—	—
90th percentile.........	61	88	116	125	139
Median	36	53	79	92	105
10th percentile.........	8	27	41	55	75

TABLE 2

DISTRIBUTION OF SCORES IN PART 4.

	ES 7 and 8	6	4 and 5	3	ES 1 and 2
No. of Cases	109	237	54	68	32
Range of Scores	% of cases	% of cases	% of cases	% of cases	% of cases
100—109	—	—	—	—	6
90—99	—	—	—	—	3
80—89	—	—	—	—	12
70—79	—	—	—	1	19
60—69	—	—	2	13	16
50—59	—	—	2	19	23
40—49	—	—	5	35	12
30—39	—	—	11	19	9
20—29	1	8	26	7	—
10—19	17	49	39	6	—
0—9	82	43	15	—	—
90th percentile.........	12	19	40	62	90
Median	5	11	18	43	68
10th percentile.........	0	4	9	22	39

As expected, there was a steady increase in the level of test scores from ES 8 to ES 1, but it will be seen that in Part I (the pictorial intelligence test) there was considerable overlap between the groups. Nearly half the members of ES 3 were below the 90th percentile for ES 6 (88). In Part 4, the general information test, there was remarkably little overlap. There was a big improvement between ES 6 and ES 4 and 5, and an even bigger improvement between ES 4 and 5 and ES 3. The 10th percentile for ES 3 (22) was above the 90th percentile for ES 6 (19).

For Parts 2 and 3 the distributions were intermediate between those for Parts 1 and 4. That is, the overlap between the different educational groups was less than for Part 1, but greater than for Part 4.

The contrast between ES 6, the normal primary school group, and ES 3, the school certificate group, is made clearer by taking ES 3 as the standard group. Let us define scores above the 90th percentile for ES 3 as 'High', between the 90th and 10th percentiles as 'Medium', and below the 10th percentile as 'Low'. The corresponding percentages for ES 6 are as follows:

Part 1	High 0, Medium 48, Low 52.
Part 3	High 0, Medium 36, Low 64.
Part 2	High 0, Medium 13, Low 87.
Part 4	High 0, Medium 6, Low 94.

The most striking contrast is between Parts 1 and 4. By ES 3 standards, 48 in every 100 of the ES 6 group have adequate intelligence, but not more than 6 out of these 48 are of adequate educational-cultural level.

For ES 1 and 2, the corresponding percentages for Part 4 are High 53, Medium 47, Low 0.

It may be objected that the tests, being short, were far from perfectly reliable. This is, of course, true but the unreliability of the separate Parts of the G.C. Test would tend to blur differences between groups, not to exaggerate them. The differences between the educational groups would have been found to be even more marked if longer, more reliable, tests had been used.

IV.—CONCLUSIONS.

We realize that a large proportion of our sample will have been affected to some extent by evacuation and other upheavals due to the war, and that this may have had a more serious effect on the primary school population than on the grammar school population, which was usually better provided for. This may have influenced our results, but we cannot say how much. The inference can, however, reasonably be drawn that there is very little likelihood of a child with normal primary school education reaching even a tolerable educational-cultural level. There is almost no overlap between children leaving school at 14 and those with education up to higher school certificate level.

Finally, we have to thank the War Office and the Civil Service Commissioners for permission to quote information from official sources, while making it clear that the opinions expressed in this paper are entirely personal.

APPENDIX

The tests included in the first experiment consisted of the Army 'General Service' battery of five tests given in the usual way, viz., Dominoes (non-verbal), Problems (pictorial test of mechanical information and aptitude), Arithmetic, Verbal and Instructions; the four parts of the general classification test; N.I.I.P. test 70/1 and a specially constructed test very similar to it; and Instructions test 'B', an alternative instructions test.

Inter-correlations were as follows :

Test	1	2	3	4	5	6	7	8	9	10	11
1. G.C. Test Part 1		2									
2. 70/1	·689		3								
3. Parallel to 70/1	·675	·727		4							
4. Dominoes	·634	·677	·622		5						
5. Problems	·588	·602	·558	·630		6					
6. G.C. Test Part 3	·606	·701	·626	·737	·636		7				
7. Arithmetic	·637	·687	·618	·764	·630	·766		8			
8. Verbal	·580	·641	·594	·675	·625	·781	·757		9		
9. Instructions	·663	·710	·641	·724	·622	·733	·796	·735		10	
10. G.C. Test Part 2	·627	·693	·632	·691	·597	·721	·741	·719	·804		11
11. Instructions B	·591	·633	·579	·658	·565	·719	·730	·752	·715	·759	
12. G.C. Test Part 4	·519	·598	·528	·589	·563	·708	·669	·753	·693	·718	·703

The correlation matrix was analysed by Miss M. S. Stevenson, to whom our thanks are due, by the method of Simple Summation, and factor-loadings were obtained as follows :

	Factors			Communality	
Test	I	II	III	Obtained	Estimated
1. G.C. Test Part 1	·760	– ·298	– ·077	·673	·670
2. 70/1	·823	– ·229	– ·091	·737	·735
3. Parallel to 70/1	·761	– ·307	– ·122	·688	·684
4. Dominoes	·825	– ·065	·183	·718	·717
5. Problems	·728	– ·090	·111	·550	·550
6. G.C. Test Part 3	·865	·066	·160	·778	·776
7. Arithmetic	·871	·059	·115	·775	·774
8. Verbal	·853	·210	·070	·777	·776
9. Instructions	·873	·049	– ·033	·769	·770
10. G.C. Test Part 2	·866	·134	– ·220	·816	·814
11. Instructions B	·826	·196	– ·051	·723	·722
12. G.C. Test Part 4	·786	·267	– ·034	·690	·689
Contribution to variance	67·5%	3·5%	1·4%		

The general factor accounted for 67·5 per cent. of the variance. Neither of the bi-polar factors was significant, but they divided the tests into easily explicable groups. The first separated the pictorial from the non-pictorial, partly educational tests, while the second separated tests of the " Code Completion " type from the remaining pictorial tests, and tests of the " Instructions " type (including general information) from the remaining non-pictorial tests.

Arranging the four parts of the general classification test according to their loadings of factor II, they rank in the order 1-3-2-4, and the factor loadings of Part 3 (·066) and Part 2 (·134) are nearer to that of Part 4 (·267) than to that of Part 1 (– ·298).

AN EXPERIMENTAL SURVEY WITH PROGRESSIVE MATRICES (1947)

By G. A. FOULDS and J. C. RAVEN
(Department of Psychological Research, The Crichton Royal, Dumfries).

I.—*Introduction.* II.—*The experimental data.* III.—*Item analysis.* IV.—*Re-test reliability.* V.—*Validity of the test results.* VI.—*Speed and efficiency of work.* VII.—*Working norms.* VIII.—*Summary and conclusions.* IX.—*References.*

I.—INTRODUCTION.

A TEST is occasionally asked for which will indicate in a few minutes whether a person is intellectually " dull," " average " or " bright." There is also a demand for reliable tests of intellectual efficiency which will differentiate clearly between people of more than average intellectual capacity. The construction of each kind of test presents difficulties. The following communication describes an attempt to overcome them.

Dull people find abstract ideas difficult to manipulate. Many verbal tests favour the verbally fluent but often fail to differentiate adequately between persons of superior intellectual capacity. For the selection of students wishing to pursue technical courses, " aptitude " and " performance " tests have been used, but for the limited information they provide, the time consumed in administering them is not always justified. *Intellectual efficiency,* in the sense of *a person's speed of accurate intellectual work,* is one essential factor determining a person's rate of progress and the degree of success in any course of training he undertakes, but depends upon a person's physical and mental health as well as upon his inborn capacity for intellectual activity.

Progressive Matrices (1938)[1] was designed to cover the whole range of intellectual development from infancy to maturity, to measure intellectual capacity independently of verbal ability or manual dexterity and to be equally useful with persons of all ages and levels of ability. In view of its wide applicability, it was expected to have certain limitations, but it was considered that from the information obtained with it,[2] it would be possible to prepare derivatives of the test designed to meet particular requirements.

Progressive Matrices (1947)[3] is a non-verbal test of intellectual efficiency designed for use with people of more than average intellectual ability. In its original form, drafted in 1943, it was used at War Office Officer Selection Boards. The 1947 revision is based upon information obtained with the 1943 scale and is designed to increase the test's reliability and range of usefulness. It is constructed to assess the ordinal, analytical and integral operations involved in higher thought processes and to differentiate clearly between people of superior intellectual ability. Over verbal tests, it has the advantage that certain thought processes are assessed independently even of elementary educational attainments. Over performance tests, it has the advantage that the information obtained is more consistent and reliable.

The scale is arranged in the form of two Sets, printed separately. Set I contains twelve problems covering the intellectual processes covered by Progressive Matrices (1938). It can be used to show whether a person may be regarded as intellectually " dull," " average " or " bright." It is also used as a practice test for Set II.

In Set II there are forty-eight problems. In presentation and argument, they are similar to those in Set I, but they increase in difficulty more steadily and become considerably more complex. They are placed in groups of four according to argument and difficulty and they are arranged so that the validity of the total score does not depend upon everyone attempting all the problems in the scale before stopping. The problems are chosen so that, at the first attempt, no one of even outstanding intellectual capacity solves them all in less than half-an-hour, while a person's interest is usually maintained for well over an hour without boredom.

II.—THE EXPERIMENTAL DATA.

To assess the consistency, re-test reliability and psychological validity of the scale, Progressive Matrices (1947), Sets I and II, were given by Mr. Foulds to the following groups of people :

471.... Children, 10½(±½) years of age, attending two Elementary Schools selected to represent all scholastic grades.
  109 Children, 10½(±½) years of age, attending two Elementary Schools, were re-tested after an interval of six weeks.
373.... Children, 12½(±½) years of age, forming the entire child population of this age at schools in a rural county town.
  111 Attending a grammar (secondary) school.
  262 Attending a central senior school and divided into six streams according to intellectual status and the course they were studying.
  92 Central senior school children, 12½(±½) years of age, were re-tested after an interval of six weeks.
223.... Children, 12½(±½) years of age, living in an urban industrial district.
  61 Attending a grammar (secondary) school.
  74 Attending a selective central school.
  88 Attending a modern (senior) school.
152.... Students, 15½(±½) years of age, taking mechanical courses at junior technical colleges.
151.... Students, 17+ years of age, taking technical and commercial courses at a Senior technical college, having obtained the school certificate or matriculation.
  243 Junior and senior technical college students were re-tested after an interval of two months.
170.... University students comprising :
  66 Arts graduates and second-year social science students.
  77 Third-year medical students.
  27 Science graduates taking the teachers' diploma.
  82 Students, after 40 minutes' work, were allowed unlimited time to finish.
304.... Adults attending a technical teachers' training college were given progressive matrices (1938) without a time limit and, some days later, Progressive Matrices (1947), Set II.
1,844 Total tested.

The subjects were told it was a test of observation and clear thinking. Set I was carefully explained to them. After a short pause, Set II was given with a time limit of forty minutes. In some cases, people were afterwards asked to go on working until they had finished all the problems in the set.

III.—ITEM ANALYSIS.

To estimate the consistency with which each problem in the scale contributed in order to the total estimate, every test protocol was used. For each group of people who obtained the same total score, the number of correct answers for each test item was recorded.*

* Fuller information and graphs are given in Progressive Matrices (1947), Plan and Use of the Scale, with the Report of an Expermental Survey by G. A. Foulds. (Obtainable from The Crichton Royal, Department of Psychological Research, Dumfries.)

In Tests 9 and 27 of Set II, the agreement between the percentage of passes and the total score is unsatisfactory. For a period at least, success in solving each of these problems decreases as the total score increases. Problem 46 presents difficulty. Most adults of superior intellectual capacity conclude that alternative No. 4 or 5 is correct, but cannot decide between them with certainty. There is insufficient data to determine how far this affects the diagnostic value of the test. The forty-eight problems constituting the scale are obviously not arranged in their absolute order of difficulty, but they are intentionally arranged in order of argument. In general, Problems 9 and 27, and possibly also Problem 46, should be re-cast. In Problems 22, 28, 30, 33, 34, 35, 36, 38, 42 and 45, minor improvements could be made in the group of alternatives from which choice has to be made. The order of presentation could also be revised. With these exceptions, the total number of problems solved in Set II appears to provide a consistent estimate of the intellectual efficiency with which a person is able to observe figures presented for his immediate perception, form comparisons between them and reason by analogy.

IV.—RE-TEST RELIABILITY.

To determine its re-test reliability, selected groups of people were given Set II to re-work after an interval of six to eight weeks. Table I summarizes the results obtained.

TABLE I

PROGRESSIVE MATRICES (1947), SET II—CORRELATION BETWEEN SCORES ON FIRST AND SECOND TESTS.

Age	No. Tested	1st Test Mean	1st Test S.D.	2nd Test Mean	2nd Test S.D.	r
10½ years ..	109	13·04	5·93	15·07	6·00	0·76
12½ years ..	92	13·77*	6·64	16·38	6·74	0·86
Adult Students	243	22·06	8·95	25·59	9·38	0·91

On repeating the test, there is a consistent increase of about three additional problems correctly solved in forty minutes. It is clear that the test cannot be used satisfactorily before the age of 11 years. At 12½ years, it is as reliable as most other tests, while with adults of more than average intellectual capacity it has the high re-test reliability of 0·91.† To determine causes contributing to the low re-test reliability of 0·76 obtained when the test was used with children under 11 years of age, the data for these children were divided into two groups according to whether scores on the first test were above or below the median for the whole group. The re-test reliability for the upper 50 per cent. was no greater than that for the lower 50 per cent. The probable explanation is that the ability to form comparisons and reason by analogy is at this age too recent an intellectual development for it to be exercised with a consistent degree of efficiency. We do not know at present how far this applies only to speed tests of intellectual efficiency and not to tests in which success depends primarily upon the utilization of acquired information, or to tests in which no time-limit is used. The reliability with which a child's intellectual efficiency can be assessed by such tests as these before the age of 11+, requires to be further investigated. It is first necessary to know how far low re-test reliability represents unreliability of the test itself, or variability of the activity assessed. We have also to determine how far, at any age, low intellectual efficiency tends also to be accompanied by increased variability of mental work.

---

* As shown in the summary of the experimental data, the group of children 12½ years of age re-tested did not contain the " brighter " children of this age.

† The reliability of a test constructed to assess a person's immediate output of intellectual work, independently of his acquired information, is always less than the maximum reliability which can be obtained with a test in which success depends upon the utilization of acquired knowledge.[4]

V.—VALIDITY OF THE TEST RESULTS.

Table II shows a comparison between the test results for selected groups of children and adults. The difference between the schools in the dockland and suburban areas was in locality rather than in teaching methods. It is interesting that in the suburban area, 20 per cent. of the children obtained scores of 21 or more compared with only 5 per cent. in the dockland area. Allowing for non-intellectual factors mediating against a child securing a scholarship to a secondary school, the test results show a striking agreement with the actual number of children taking up scholarships from these two areas.

TABLE II

SET II—SCORE DISTRIBUTION FOR SELECTED GROUPS OF CHILDREN AND ADULTS.

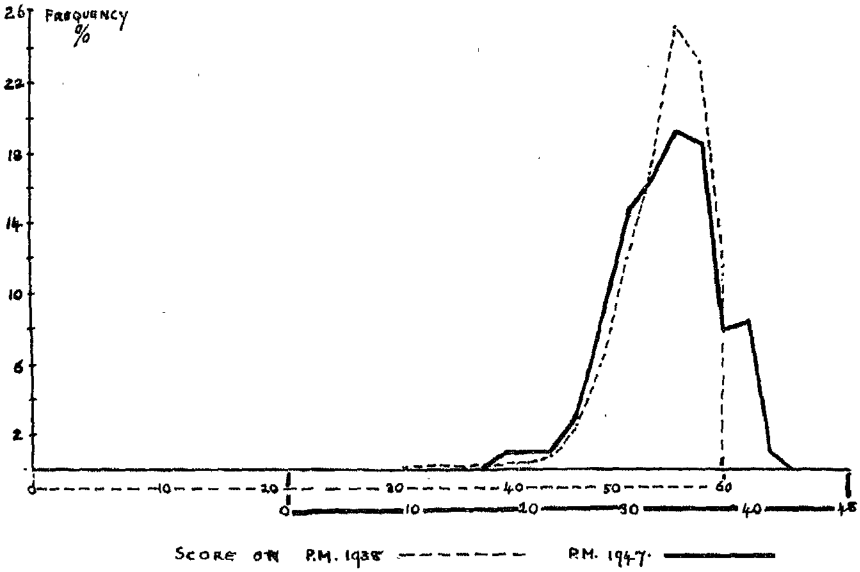
Urban		No.	Mean	S.D.
A. Elementary School Children 10⅞ ± ½ years	Dockland Area ..	129	10·3	5·1
	Suburban Area ..	222	14·5	5·9
B. Urban Children 12½ ± ½ years attending	Senior Modern School	88	12·5	6·3
	Selective Central School	74	22·2	4·5
	Grammar School ..	61	25·6	6·3
C. Rural Children, 12½ ± ½years, attending	Central School Course F	40	8·2	4·1
	,, ,, E	40	14·9	5·1
	,, ,, D	42	14·5	3·8
	,, ,, C	36	14·6	4·2
	,, ,, B	40	15·6	4·7
	,, ,, A	64	20·0	4·8
	Grammar School ..	111	24·0	4·9
D. Junior Technical College Students (Mechanics) ..		152	24·9	6·1
Senior Technical and Commercial College Students		151	29·8	5·1
University Students	Faculty of Arts ..	66	32·1	3·5
	Faculty of Medicine ..	77	33·1	4·2
	Faculty of Science ..	27	34·1	3·2

For children 12½ years of age living in an urban area, and attending the senior modern school, the selective central school, and the grammar school in the same city, the mean scores on the test are significantly different. (The selective central school caters for those children who have just failed to obtain scholarships.)

The score distributions for children 12½ years of age, living in the rural area, and attending the central senior and grammar school for the district, are also interesting. The central senior school caters for all the children who do not pass the examination for the grammar school, and they are placed in six educational streams in ascending order from " F " to " A," according to the child's ability and the type of course he is pursuing.

Table II also shows a comparison between the test scores obtained by junior and senior technical college and university arts, science and medical, students. The number of university students in each Faculty is too small for definite conclusions to be drawn from the results. The small number of Science students in the sample may account for the dip occurring in Graph I, a little above the mean.

GRAPH 1



In Graph I, the score distribution on Progressive Matrices (1947), Set II, for the whole group of 170 university students, is compared with the untimed Progressive Matrices (1938) score distribution for 881 students forming a representative cross-section of university students tested by Dr. Leybourne-White. On Progressive Matrices (1938) the scores obtained by university students cluster a-symmetrically at the upper end of the scale, while approximately two-thirds of the items fail to differentiate between them. On Progressive Matrices (1947), Set II, the scores obtained by University students are distributed more symmetrically about the mean score of 32·9 and a standard deviation of 3·9. Even for university students, the scale gives an effective range nearly four times the standard deviation above the mean, and there is no evidence of any ceiling effect either at the end or at any intermediate point in the scale. Above the mean, there is one small dip in the curve which has already been accounted for. Below the mean, there is a small secondary nodal point, but this has occurred in other investigations with university students. It appears to represent students who have entered the university on other than scholastic grounds. From the results obtained, perceptual tests of eduction appear quite suitable for eliciting the higher intellectual functions and for assessing superior intellectual efficiency.

VI.—SPEED AND EFFICIENCY OF WORK.

At the end of forty minutes' work at speed, selected groups of students were asked to go on working until they had completed all the problems in Set II. The relationship between the scores obtained in limited and unlimited time is shown in Table III.

Working at speed, some students attempt a large number of problems, apparently being prepared to guess when a solution is not clear to them. Other students attempt fewer problems, but make fewer mistakes. They apparently make certain that they have solved each problem correctly before going on to the next. In either case, from Table III, the output of efficient intellectual activity appears to be almost the same. The chief difference is that the former type of worker uses up more of his test material in forty minutes, and as a result cannot greatly increase his score when he is afterwards told to go on working until he has finished the set. From Table III, it also appears that the maximum output of efficient intellectual work is a judicious utilization of both material and time, resulting in a kind of cautious guessing intermediate between the two extremes.

TABLE III

RELATION BETWEEN SPEED AND SCORE ON SET II.

Total Time in Minutes to Finish	No. of Students	In 40 minutes.		In Unlimited Time	
		Mean Score	Mean position of problem being worked on	Mean Score	Increase after 40 minutes
40–54	14	32·9	45th	33·4	0·5
55–69	33	34·4	41st	37·4	3·0
70–84	20	32·8	38th	37·3	4·5
85–99	15	32·7	37th	38·2	5·5

The amount of test material used up in the forty minutes, during which time the students did not know they would be allowed additional time to finish, naturally affected the number of problems available for solution later. The total scores obtained in unlimited time are, therefore, not identical with the scores students might have obtained if, from the outset, they had been given the series as an untimed test. Students who, in 40 minutes, attempted all but the last four problems in the scale, were seriously handicapped. Those who had not attempted more than forty problems tended to obtain the same mean maximum scores provided they were allowed sufficient time. In other words, the students differed more in their rate of work than in their intellectual capacity. Performance in these tests appears to depend to some extent upon the economical utilization of the time at one's disposal, but the scale can apparently be used satisfactorily to assess intellectual efficiency.

VII.—WORKING NORMS.

The total sample of 471 children at 10½ years of age, and 596 children at 12½ years of age, was used to establish working norms for the test. By comparing the test scores of 304 adults given Progressive Matrices (1938), as an untimed test, and a few days later given Progressive Matrices (1947), Set II, as a forty-minutes' efficiency test, it was also possible to establish working norms for adults of 20 years of age. Table IV shows for Progressive Matrices (1938), the fiftieth, seventy-fifth, ninetieth and ninety-fifth percentile scores at 10½, 12½ and 20 years of age, and the estimated percentile scores at the same ages for Progressive Matrices (1947), Set II. As the latter test is designed primarily for use with persons of average or more than average intellectual ability, percentile scores below the median are not shown. For scores lying between 20 and 30 there is an approximately one-to-one agreement with scores obtained with Progressive Matrices (1938). For scores greater than 30, Progressive Matrices (1947), Set II, appears to provide an increasingly accurate differential estimate of intellectual efficiency.

TABLE IV

WORKING PERCENTILE NORMS.

Matrices Test Used		1938, with no time limit			1947, Set II, with a time limit of 40 minutes		
Age		10½	12½	20	10½	12½	20
No. of persons tested		1,407 children and 3,665 men			471	596	304
Percentile Points	95	48	52	55	23	30	34
	90	45	50	54	21	28	31
	75	39	46	49	17	24	26
	50	33	41	44	12	19	21

For groups of " dull," " average " or " bright " persons, a comparison was made between the test scores on Progressive Matrices (1938) and on Set I of Progressive Matrices (1947). In ten minutes, the " dullest " 10 per cent. of the population seldom solved more than the first five consecutive problems of Set I, although they occasionally got one additional problem right, more or less by chance. Conversely, in ten minutes, the intellectually " brightest " 10 per cent. of the population rarely made more than one mistake. As a quick ten-minute test, less than six consecutive solutions in Set I, followed by consistent failure, can apparently be accepted as indicating intellectual " dullness ," while one or no mistakes can be accepted as indicating that a person is intellectually " bright."

VIII.—SUMMARY AND CONCLUSIONS.

From the results of the survey the scale appears to provide a consistent estimate of the efficiency with which a person is able to apprehend characters presented for his perception, form comparisons between them and reason by analogy. Low re-test reliability before the age of eleven years does not appear to be due to the design of the test alone, and opens a question of considerable psychological interest. The scale's re-test reliability increases rapidly after the age of eleven years. Used with adults, the scale appears to have a high re-test reliability, while the results obtained on the test agree with independent estimates of intellectual ability. The test scores obtained by university students in forty minutes show no evidence of ceiling effects either at the end of the scale, or at any intermediary level.

We are indebted to the many people who voluntarily co-operated in the survey, and to Mr. M. H. Paterson, Mr. G. Tweddle and Miss E. A. White, for their assistance in tabulating the data obtained. Thanks are also due to Dr. P. K. McCowan and the Directors of The Crichton Royal for making the necessary research funds available for the work.

IX.—REFERENCES.

[1] RAVEN, J. C.: *Progressive Matrices*, 1938 (H. K. Lewis and Co., Ltd., London).

[2] RAVEN, J. C.: " The R.E.C.I. Series of Perceptual Tests : An Experimental Survey " (*Brit. J. Med. Psych.*, Vol. XVIII, Part I, 1939).

[3] FOULDS, G. A., and RAVEN, J. C.: *Progressive Matrices* (1947), *Sets I and II* (For Adults) (H. K. Lewis and Co., Ltd., London).

[4] RAVEN, J. C.: " The Comparative Assessment of Intellectual Ability " (*Brit. J. of Psych.*, Vol. XXXIX, Part I, 1948).

Typescript received May, 1949.

A NEW TEST OF POETRY DISCRIMINATION.*

By E. M. EPPEL

I.—*Introduction.* II.—*Aims of investigation.* III.—*Construction of the test.* IV.—*Method.* V.—*Example of test.* VI.—*Groups tested.* VII.—*Conclusions.* VIII.—*General educational implications.* IX.—*References.*

I.—INTRODUCTION.

It is surprising that in this country, with its exceptionally high standard of poetic achievement and its vast body of critical material on execution and judgment, so few attempts have been made at controlled investigation of levels of appreciation or of accompanying mental processes. Richards[1] made an important qualitative study of the responses of a group of students of English Literature, but the implications of his work were limited and had the weaknesses of all subjective approaches. Williams, Winter and Woods[2] carried out a series of tests of æsthetic appreciation on a group of school children. The poetry tests which were included broke valuable new ground. One test, however, used so unsatisfactory a poetic unit as the couplet, and involved comparing "earlier" and "later" versions by the same poet. The "later" (usually accepted as the "better") version would often be well-known to school children, in this country, e.g., "La Belle Dame Sans Merci" and "We Are Seven." Another test involved grading three extracts from poems on the same topic. The *a priori* grading against which to measure the subjects' performance was based on such criteria as inclusion of poems in *Oxford Book of English Verse* for the best, and in popular magazines for the worst. These categories are useful so far as central tendencies are concerned, but the border-line areas may raise difficulties. This way does not seem to be the best technique for the comparison of one poet with another.

In the U.S.A., Abbott and Trabue[3] constructed a test in which original versions were presented along with versions altered in what they considered to be essential characteristics of poetry. These were often too radically altered, however, and also many of the poems used would be known to large numbers of the subjects. It is not surprising that English specialists scored highest on this test. Vergera[4] made an investigation to secure qualitative responses of students to the sense and mood of poems. Her approach involved reading poems under a variety of controlled conditions, and the use of case histories. Unfortunately, she dealt with a comparatively small number of students, only women.

II.—AIMS OF THE PRESENT INVESTIGATION.

The aims of the present investigation were :

(1) The creation of an adequate test of differences in the ability to discriminate poetical quality.

(2) The assessment by means of this test of any changes or developments in this ability, which occur between the sexes and age groups of senior secondary school and university populations.

---

* Part of a Symposium on New Objective Tests of Secondary School Subjects, British Psychological Society, Scottish Branch, 26th November, 1949. (Based on a thesis approved as part-requirement for the Degree of Bachelor of Education, at Glasgow University, 1949.)

(3) The examination of sex and age preferences for :
  (*a*) Two clearly defined types of poetry, designated " Direct " and " Oblique."*
  (*b*) Poetry making use of two clearly defined kinds of subject matter, designated " Nature " and " Human Relationships."

III.—CONSTRUCTION OF THE TEST.

In this test a number of poetic extracts were presented with a line or lines missing. The criterion of discrimination was the selection, to fill the gap, of the best, i.e., the original, line from among a number of artificially weakened variants. The weakening process was carried out by a careful analysis of the contribution of the line to the texture of the poem, and by a deliberate attempt to break down those elements of sound, imagery, rhythm and meaning, which gave the line its peculiar quality and integrated it with its context of preceding and succeeding lines. This method differed from others previously used in that it took account of the value of the poetical elements of the line relative to the homogeneity and unity of the whole extract. It did not involve the comparison of completely different poems, or of different authors. Ideally, the whole poem should have been printed in each case to act as context and poetic unit, but a compromise had to be effected between the ideal and the practical, and no extract contained less than four lines, while all were reasonably complete in themselves.

It was considered that the following principles were essential to the test of poetic discrimination :

(1) That the poems should be of high literary merit.
(2) That the poems should represent a fair range of styles and periods.
(3) That the poems should, as far as possible, be unknown to the subjects, and should certainly not include any taught in the secondary schools from which most of the subjects were drawn.
(4) That the vocabulary on the *denotative* side should not be so difficult as to inhibit or penalise unduly the youngest subjects to be tested.
(5) That every precaution be taken to ensure that no variation could be considered by competent judges to be as good as, or better than, the original.
(6) That the only considerations for the detection of the correct line should be of an æsthetic nature, e.g., since a highly educated section of the community was being tested, it was necessary to provide variations which were not too grossly prosaic. To avoid undue bias being given to intellectual detection of incongruities of meaning, each of the weaker variations had to make reasonable sense in the passage. Any rhyming words had to appear in all the versions, and only very subtle changes in metre were allowed, to obviate the possibility of subjects applying mechanical rules for the detection of the more common prosodic measures.

It was necessary to elicit a consensus of expert opinion to establish most of these criteria, as it was to establish into which of the two types of poetry—" Direct " or " Oblique "—each extract fell.

---

* These terms are derived from E. M. Tillyard[5], and are familiar to specialists in English Literature. Briefly, the poetry of Direct Statement contains few, if any, references beyond those explicitly stated ; the poetry of Oblique Statement is highly evocative and may imply emotions or ideas, etc., not explicitly stated.

IV.—Method.

A list of seventy-eight poems, ranging from the fifteenth to the twentieth century, was presented to a group of experienced English teachers from various schools in the Glasgow area. Any poem read at any stage in the school was eliminated from consideration as a possible test item. Extracts were made from forty-six of the remaining poems. A line or couplet was taken from each and presented beside the context, with weakened variations. This preliminary test was presented to a group of experts (eleven females, fourteen males). Most were teachers of English Literature in schools or universities ; some were poets ; all had a wide knowledge and experience of poetry. They were asked to evaluate each extract on the following counts :

(1) Whether each variation was inferior to the original or not.

(2) Whether the extracts were " Direct " or " Oblique."

(3) Whether any item was not good enough to be included in a test of poetry discrimination.

Tabulation of results and statistical verification of significant differences enabled ten " Direct " and ten " Oblique " items to be chosen to compose the final test. These were arranged alternately and without indication of differences. Of the items chosen, four clearly dealt with Nature, and four with Human Relationships. In each case two of the items were " Direct " and two " Oblique."

Information on the tests were phrased simply and clearly, and all responses were made on a separate score sheet by ringing (*a*), (*b*) or (*c*) for the variation chosen. Preferences were also indicated by ringing l (like), i (indifferent), d (dislike), for each item. In this investigation only the positive preferences were analysed.

The information to be extracted from each score sheet was :

Total Score	20	Liking, Oblique Items	10
Score on Direct Items	10	Liking, Nature Items	4
Score on Oblique Items	10	Liking, Human Relationship Items	4
Liking, Direct Items	10		

An example using a well-known poem was done in the test, and the subject was asked to study this before starting to deal with Item No. 1.

No time limit was intended, but in practice forty minutes was found to be adequate for almost all subjects. Standard instructions to testers were issued, and by the use of suitable templates, it was possible to score and extract data in two rapid operations.

The following is a reproduction of the instructions and the first item of the test :

V.—Example of test.

POETRY DISCRIMINATION

Read these instructions carefully.

In each of the poems printed below, there is a line missing, indicated by asterisks (* * *). Below each poem are printed three possible versions of the missing line. Any of these versions could go into the poem and make sense, but one is the original and is considered to be the best poetically.

Try to imagine how these poems would sound if read aloud, and decide which line, *a*, *b*, or *c* is the one which fits in best. When you have done this put a ring round the corresponding letter on the score-sheet (see the example which has been done for you).

You are asked also to say whether you *like* the poem ; are *indifferent* to it (that is, don't care much one way or the other) ; or *dislike* it. Do this for each item by putting a ring, on the score sheet, round *l* if you like it, *i* if you are indifferent, or *d* if you dislike it (see the example).

After you have studied the example, start with Item No. 1. Don't hurry, but work steadily, and try to do all the items.

Example.

I wandered lonely as a cloud

* * * * * * *

When at all once I saw a crowd,
A host of golden daffodils !

(*a*) That moves quite slow past woods and hills,
(*b*) That floats on high o'er vale and hills.
(*c*) That glides above flat plains and hills.

In this example, (*b*) is the original line, so there is a ring round (*b*) on the score-sheet. The person who made up this test *likes* this poem so he has put a ring round l on the score-sheet.

Item No.

1. O Earth lie heavily upon her eyes ;

* * * * * * * *

Lie close around her ; leave no room for mirth

With its harsh laughter, nor for sound of sighs . . .

(*a*) Close her pale lids, weary of looking, Earth ;
(*b*) Shut her dear eyes tired of watching, Earth ;
(*c*) Seal her sweet eyes weary of watching, Earth ;

VI.—GROUPS TESTED.

Six-hundred subjects were tested comprising five male and five female groups of sixty each. For both sexes the groups were as follow :

(1) 13-year-olds ; (2) 15-year-olds ; (3) 17-year-olds ; (4) undergraduates ; (5) graduates.

For age groups 13-17 years two senior secondary schools with similar socio-economic backgrounds were used. The total age group in the first school was used, and numbers made up by random selection from the second school. The university subjects comprised students, research workers, and staff of two large, non-residential universities. The average age of adult subjects was as follows :

	Male	Female
Undergraduates . .	23·7 years	22·3 years
Graduates	26·6 years	30·2 years

Satisfactory reliability coefficients were calculated for the test by the "split half" method. This was done on two randomly chosen male and female groups of fifty each, and one complete female age group (15 years).

Indices of discrimination were calculated for each item on male and female groups separately ; all were positive. Final I.D.'s were arrived at by averaging male and female I.D.'s for total groups. Seventeen of the twenty items had I.D.'s ranging from + ·5 to + ·85 and none had less than + ·3.

The distributions of scores on the total test for all groups, both male and female, indicate that the majority of subjects tend to score around the average, and that there is a steady diminution in frequency towards the highest and lowest scores.

VII.—CONCLUSIONS.

The technique adopted seems to be a fruitful one. The test is easily administered, arouses interest, and requires a minimum of explanation.

The following main conclusions may be drawn :

1.—It was found that for both sexes there is a steady and highly significant increase in total score with increasing age, which suggests that poetic discrimination develops consistently throughout and beyond adolescence.

2.—The highly significant differences in ability between 17-year-old and graduate groups may be due to the further selection involved (including higher ' g ' and more intellectual interests). Poetry discrimination, however, may be an ability susceptible of even further training, and the extension of cultural background.

3.—It was found that at most group levels the female subjects scored higher than the males. The summation of the tendencies was significant at a level between 5 per cent. and 2 per cent. As yet it is not known whether this discrepancy should be explained by reference to innate sex differences, or to differences in educational and cultural influences.

4.—A rather unexpected result was that all groups of both sexes scored significantly higher on the Oblique items than on the Direct. Since there are no substantial differences in the Indices of Discrimination of these two types of poetic extracts, the reason would appear to lie in the intrinsic differences between the two types of poetry, and in their respective histories as art forms and teaching material.

5.—Interesting and convincing results were obtained in the examination of group preferences (liking) for Direct and Oblique items. The trends were the same for both sexes, but more clearly significant for the females.

(*a*) In earlier adolescence (13-15 year groups) there is a decided preference for the poetry of direct statement.

(*b*) In later adolescence greater catholicity of taste is shown. In fact, no significant differences of preference for Direct or Oblique items were found at 17 years.

(*c*) The adult groups showed a preference for Oblique items as compared with Direct. This was distinct and highly significant in the case of female Undergraduate and Graduate groups. ($p = \cdot 001$ in both cases.)

6.—No significant conclusions could be drawn from the investigation of preferences for poems using different kinds of subject matter. Certain tendencies for a shift of preference from ' Nature ' to ' Relationship ' items were observed in both sexes with increasing age, but it is felt that the items in each category were too few for reliable results. The problem seems worth investigating with a larger number of items specially chosen for the purpose.

VIII.—GENERAL EDUCATIONAL IMPLICATIONS.

The comparatively low average scores of the oldest and most highly educated groups tested, seems to bear out Richards' observation that the general level of critical appreciation in such groups is not of a high standard. Training in the analysis of specifically poetic elements may be a valuable aid to critical discrimination. The total impression of first acquaintance should be followed by an examination of those elements of sound, rhythm, imagery, and meaning that go to the making of that impression. Even at higher levels much seems to be taken for granted about the student's knowledge of the components of poetry. The analysis must, of course, be incorporated into and followed by a new syntheses, the fresh total impression, which will provide a richer and more complex experience for the reader, and also raise his standards of critical discernment in his approach to other poems.

IX.—REFERENCES.

[1] RICHARDS, I. A.: *Practical Criticism: A Study of Literary Judgment* (Kegan Paul, 1929).

[2] WILLIAMS, E.D., WINTER, L. and WOODS, J.M.: " Tests of Literary Appreciation " (*Brit. J. Educ. Psych.*, 1938, Vol. VIII, Part 3, pp. 265-285).

[3] ABBOTT and TRABUE: " A Measure of Ability to Judge Poetry " (*Teachers' College Record*, Vol. XXII, No. 2, 1921, pp. 101-126).

[4] VERGERA, A. D.: *A Criticial Study of a Group of College Women's Responses to Poetry* (Bureau of Publications, Teachers' College, Columbia University, New York, 1946).

[5] TILLYARD, E. M.: *Poetry, Direct and Oblique* (Chatto and Windus, 1934).

ATTITUDE TOWARDS WELSH AS A SECOND LANGUAGE

A FURTHER INVESTIGATION.

BY W. R. JONES
(*Lecturer in the Department of Education, University College, Bangor*).

I.—*Introduction.* II.—*Scales for measuring* (*a*) *attitude towards Welsh, and* (*b*) *linguistic background.* III.—*Application of the scales.* IV.—*Tabulation of the results.* V.—*Discussion of results.* VI.—*The relationship between attitude and intelligence.* VII.—*Correlation between attitude and attainment.* VIII.—*Summary and conclusions.*

I.—INTRODUCTION.

IN a previous article[1] the writer described a preliminary investigation of pupils' attitude towards Welsh as a second language in the school curriculum. A detailed account was given of the construction of an Attitude Scale, and of its subsequent application to complete age-groups of children in a secondary modern school in the Rhondda Valley. Certain significant differences were found between various sub-groups in the enquiry, for example, between boys and girls, and between groups of children from homes of diverse linguistic structure. The purpose of the present investigation was to seek further information about pupils' attitude towards Welsh, and to discover whether the differences and tendencies previously observed are equally operative in the case of children drawn from a suburban rather than an industrial area.

II.—THE SCALES FOR MEASURING (A) ATTITUDE TOWARDS WELSH, AND (B) LINGUISTIC BACKGROUND.

(*a*) In the present investigation the general method of constructing the Attitude Scale was identical with the one employed in the preliminary enquiry—the same fifty-five original statements, expressing varying degrees of attitude towards Welsh were given to a number of judges for classification into eleven groups in accordance with the degree of unfavourable or favourable attitude which they believed were expressed in them. The frequencies of their judgments were tabulated and, thereafter cumulatively added, each cumulative frequency being calculated as a percentage of the total number of judgments ; from these cumulative frequencies, graphically represented, two values were determined for each statement—*a scale value* and *a measure of ambiguity*.[2] An attempt, however, was made to improve the validity[3] of the scale in one important respect ; whereas in the preliminary investigation only thirty-five persons took part in the classification of statements, in the present enquiry this number was increased to one hundred. In fourteen out of the fifty-five statements there was no change at all in scale value as between the first and second classification ; in the remaining forty-one statements the change was very slight indeed : twenty-six statements showed an increase in value ranging from +0·1 to +0·7, the average increase being +0·29; fifteen statements showed a decrease in scale value ranging from −0·1 to −0·4, the average decrease being −0·19. Thus, a close agreement was found between the results of the two classifications as regards the scale value of each statement. It was also found that the two classifications did not differ markedly in respect of ambiguity or variability as measured by the semi-interquartile range: the average Q values were 0·894 and 0·917 for the first and the second classifications respectively.[1]

---

[1] " Attitude towards Welsh as a Second Language—A Preliminary Investigation," this *Journal*, February, 1949, Vol. XIX, Part I, pp. 44-52.

[2] For a full explanation of the procedure involved in the construction of the Attitude Scale the reader is referred to the previous paper : *op. cit.* Appendix, Notes 1, 2, and 3, pp. 50-52.

[3] The validity of a test is normally defined as its trueness as a measure of what it purports to measure. A useful discussion of the validity of Attitude Tests with appropriate references may be found in F. J. SCHONELL'S " The Development of Educational Research in Great Britain, Part III," this *Journal*, Vol. XVIII, Part II, June, 1948, pp. 53-58.

As in the preliminary investigation, twenty-two statements selected mainly on the basis of the scale values and the Q deviations were included in the final form of the Attitude Scale. These statements were almost equally spaced along the whole scale, two statements being assigned to each one of the eleven groups of classification. In giving the scale to the pupils the statements were for obvious reasons arranged in random order.

The instructions printed clearly at the top of the test-sheet requested the pupils to read the statements carefully, and then to put a tick if they *fully* agreed with a statement, and a cross if they did *not* agree. For scoring, the scale values were used, the attitude score of any individual being the mean scale value of all statements answered in the affirmative.

The test-sheets, in addition to the attitude statements, also contained material whereby some useful information could be obtained on matters appertaining to the acquisition of a second language. It was accordingly divided into five sections as follows: Sections A and B consisted of two alternative forms of the Attitude Scale proper, and the pupils' answers to these sections alone determined their attitude score. Section C consisted of a brief list of further attitude statements, which were included for their intrinsic merit and interest. Section D was designed to determine pupils' preferences in relation to various aspects of language-learning as exemplified in the case of Welsh as a second language. For this purpose a number of sentences were set out, each one of them referring specifically either to one or to more than one aspect of language-learning.[2] Thus, the oral and passive aspect was represented by the following sentence: "I only want to learn enough Welsh to understand others speaking it"; other aspects were likewise stated in simple sentences of this kind. Pupils were then instructed to put a tick after the *one* sentence which was most true about them. In Section E of the test-sheet pupils were requested to write briefly stating why they liked Welsh or why they did not like Welsh.

(*b*) The Language Questionnaire used in the preliminary investigation was a valuable means of gaining a general impression only of the pupils' linguistic background and of classifying them accordingly into three or four broad linguistic groups. The simplicity and brevity of its design precluded accurate and detailed estimation. In order to gauge the child's linguistic experience outside the classroom more finely and accurately, a Language Questionnaire was specially constructed—a shortened and modified form of similar questionnaires which have proved useful in other countries.[1] It consisted of twelve items covering the active and passive uses of the oral form of Welsh in the home and in the playground. The following example will suffice to illustrate the nature of the questions asked :

---

[1] The product moment correlation between the two sets of Scale Values was +0·998 ±0·0004, and between the two sets of Q values +0·972 ±0·0052. These figures indicate a remarkably close agreement between the two classifications by thirty-five and one hundred judges respectively, and suggest that the use of a small number of judges (thirty-five in the present instance) may be sufficient for standardising Thurstone-type attitude tests. For a similar suggestion, cf. GLASSEY, WILLIAM: "The Attitude of Grammar School Pupils and their Parents to Education, Religion and Sport" (M.A. Thesis, University of Manchester), reported in an abstract in this *Journal*, Vol. XV, Part II, June, 1945, p. 101.

[2] Cf. PALMER, HAROLD E.: *The Scientific Study and Teaching of Languages* (George G. Harrap and Co., Ltd., 1937), pp. 225-228 on the four aspects of a given language, that is, the active and passive uses of both the oral and written forms. Various combinations of these four aspects were represented in Section D of the test.

" Answer the following questions by drawing a line under ONE ONLY of the five words to be found after each question :

Question 1 (*a*) Does your FATHER speak to you in WELSH ?
(1) Never ; (2) Sometimes ; (3) Often ; (4) Mostly ; (5) Always.

Question 2 (*a*) Do you speak in WELSH to your FATHER ?
(1) Never ; (2) Sometimes ; (3) Often ; (4) Mostly ; (5) Always.

Question 3 (*a*) Does your FATHER speak to you in ENGLISH ?
(1) Never ; (2) Sometimes ; (3) Often ; (4) Mostly ; (5) Always.

Question 4 (*a*) Do you speak in ENGLISH to your FATHER ?
(1) Never ; (2) Sometimes ; (3) Often ; (4) Mostly ; (5) Always."

A supplementary question was as follows :

Question 7 (*a*) Is your FATHER able to speak WELSH ?
(1) Yes ; (2) No ; (3) A little.
(*b*) Is your MOTHER able to speak WELSH ?
(1) Yes ; (2) No ; (3) A little.

The questions were answered one at a time after the examiner had read them aloud. Pupils were constantly reminded of the exact meaning to be attached to each one of the five words after each question ; thus, with reference to any activity mentioned in a question, it was emphasized that " never " was to be underlined if the person stated *never* did it, " sometimes," if he did it *less than half of the time*, " often " if he did it *about half of the time*, " mostly " if he did it *more than half of the time*, and " always " if he did it *the whole time*. It was also stressed that if " always " was underlined for Welsh, " never " had to be underlined for English, otherwise the evidence would be self-contradictory. The importance of answering every item was emphasized, and the papers were carefully checked to ensure that there were no omissions. Some omissions were inevitable, for example, in cases where a child's parent was dead. For scoring the following weights were given to the various items underlined : Never=0, Sometimes=1, Often=2, Mostly=3, Always=4. In this way the child's linguistic background was measured on a scale of twenty-four units.

III.—APPLICATION OF THE SCALES.

The investigation was carried out in a secondary modern school in a populous suburban area on the outskirts of Cardiff, with the full co-operation of the headmaster and the specialist teacher of Welsh. With the exception of a few absentees, every pupil who took Welsh as a second language was tested by the writer himself. The total number was 211 boys and girls, whose ages ranged from 11 : 6 to 15 : 5, grouped in eight different forms constituting the A and B streams in the school. Forms in the C stream were not included because they were not taught Welsh *as a language* ; the headmaster was of the opinion that the acquisition of a second language was beyond the limited mental powers of these pupils, and, consequently, instead of the usual lessons in Welsh, they were given a general course of instruction through the medium of English on the history and institutions of Wales. All of the pupils in the A and B streams were taught Welsh as a second language by a specialist teacher responsible for the teaching of the subject throughout the school. The Direct Method was employed, and the standard of teaching was extraordinarily high.

---

[1] Cf. HOFFMAN, M. N. H. : *The Measurement of Bilingual Background* (Teachers' College, Columbia University Contributions to Education, No. 623, Bureau of Publications, Teachers' College, Columbia University, New York City, 1934). Also MALHERBE, E. G. : *The Bilingual School—A Study of Bilingualism in South Africa* (Longmans, Green and Co., 1946).

The results obtained in answer to the Language Questionnaire may be briefly summarised: pupils' answers were first converted into scores in accordance with the 24-unit scale of marking already described, and for the sake of convenience of treatment these scores were then grouped into seven categories, (*a*), (*b*), (*c*), and so on, as shown in the accompanying table :

SUMMARY OF LANGUAGE QUESTIONNAIRE DATA

Category	Score		Number of Pupils	Percentage Pupils
	Welsh	English		
(*a*) Monoglot English	0	24	115	61·50
(*b*) Monoglot English with slight Welsh	1–5	23–19	71	37·97
(*c*) Bilingual (English)	6–8	18–16	0	0·00
(*d*) Bilingual (50 – 50)	9–15	15–9	0	0·00
(*e*) Bilingual (Welsh)	16–18	8–6	1	0·53
(*f*) Monoglot Welsh with slight English	19–23	5–1	0	0·00
(*g*) Monoglot Welsh	24	0	0	0·00
	Total Number* ..		187	100·00

* Returns which contained unavoidable omissions were not included.

It is evident that from the point of view of linguistic background, the present group is composed mainly of monoglot English-speaking children, who have no experience whatsoever of Welsh outside the walls of the classroom. Within the group, however, there are a number of pupils, who either hear or both hear and speak Welsh to some extent in their out-of-school environment, that is, monoglot English-speaking children with a slight knowledge or experience of Welsh. At first sight the entire group seems to possess the same linguistic pattern as the one previously tested in the Rhondda Valley.[1] Upon closer analysis, however, it appears to be definitely less native and indigenous to Wales than the Rhondda group, which, though outwardly anglicized, was fundamentally Welsh, and essentially, even though remotely, bound up with Welsh tradition and culture.

IV.—TABULATION OF RESULTS.

A.—*The Reliability of the Attitude Scale.*

The product moment correlation between the alternative forms (A and B) of the Attitude Scale was as follows : $r_{AB} = +0{\cdot}69 \pm 0{\cdot}024$.

Use of the Spearman-Brown prophecy-formula gave the following figure :

Reliability Coefficient $= 0{\cdot}82$.

B.—Tables IA and IB to VA and VB below show the mean attitude scores, and the difference in means and in variability between various sub-groups in the enquiry. The higher the score the more favourable the attitude indicated.

[1] JONES, W. R. : *Op. cit.*, pp. 45-46.

TABLE 1A

MEAN ATTITUDE SCORES FOR VARIOUS AGE-GROUPS

Age-groups	No.	Mean Score	S.D.	σ_m	σ_σ
I. 11 : 6—12 : 5	53	6·74	0·92	0·13	0·09
II. 12 : 6—13 : 5	61	6·64	1·38	0·18	0·13
III. 13 : 6—14 : 5	53	6·41	1·30	0·18	0·13
IV. 14 : 6—15 : 5	44	5·52	1·70	0·26	0·18

TABLE IB

DIFFERENCES IN MEAN ATTITUDE SCORES AND S.D.'S FOR VARIOUS AGE-GROUPS

Age-Groups	Mean Score	Diff. in Score	σ_{m_d}	' t '	S.D.	Diff. in S.D.	σ_{σ_d}	' t '
I. 11 : 6–12 : 5 II. 12 : 6–13 : 5	6·74 6·64	0·10	0·22	0·45	0·92 1·38	0·46	0·16	2·88 *
I. 11 : 6–12 : 5 III. 13 : 6–14 : 5	6·74 6·41	0·33	0·22	1·50	0·92 1·30	0·38	0·16	2·38 †
I. 11 : 6–12 : 5 IV. 14 : 6–15 : 5	6·74 5·52	1·22	0·29	4·21 *	0·92 1·70	0·78	0·20	3·90 *
II. 12 : 6–13 : 5 III. 13 : 6–14 - 5	6·64 6·41	0·23	0·25	0·92	1·38 1·30	0·08	0·18	0·44
II. 12 : 6–13 : 5 IV. 14 : 6–15 : 5	6.64 5·52	1·12	0·31	3·61 *	1·38 1·70	0-32	0·22	1·45
III. 13 : 6–14 : 5 IV. 14 : 6–15 : 5	6·41 5·52	0·89	0·32	2·78 *	1·30 1·70	0·40	0·22	1·82

* Significant at the 1 per cent. level.
† Significant at the 5 per cent. level.

TABLE IIA

MEAN ATTITUDE SCORES FOR " A " AND " B " FORMS

Forms.	No.	Average Age	Mean Score	S.D.	σ_m	σ_σ
IA, IIA, IIIA, IVA . .	107	13 : 5	6·27	1·45	0·14	0·10
IB, IIB, IIIB, IVB . .	104	13 : 5	6·48	1·37	0·14	0·10

TABLE IIB

DIFFERENCES IN MEAN ATTITUDE SCORES AND S.D.'S FOR " A " AND " B " FORMS

Forms.	Mean Score	Diff. in Score	σ_{m_d}	' t '	S.D.	Diff. in S.D.	σ_{σ_d}	' t '
IA,IIA,IIIA,IVA	6·27	0·21	0·20	1·05	1·45	0·08	0·14	0·57
IB,IIB,IIIB,IVB	6·48				1·37			

TABLE IIIA

Mean Attitude Scores for Girls and Boys Separately

Group	No.	Average Age	Mean Score	S.D.	σ_m	σ_σ
Girls	112	13 : 5	6·68	1·33	0·13	0·09
Boys	99	13 : 3	6·05	1·43	0·14	0·10

TABLE IIIB

Differences in Mean Attitude Scores and S.D.'s for Girls and Boys

Group.	Mean Score	Diff. in Score	σ_{m_d}	't'	S.D.	Diff. in S.D.	σ_{σ_d}	't'
Girls	6·68	0·63	0·19	3·32 *	1·33	0·10	0·14	0·71
Boys	6·05				1·43			

* Significant at the 1 per cent. level.

TABLE IVA

Mean Attitude Score for Group A=Pupils of Welsh-Speaking Parents, Group B=Pupils Whose Parents Know a Little Welsh, and Group C=Pupils of Non-Welsh Speaking Parents

Group	No.	Average Age	Mean Score	S.D.	σ_m	σ_σ
A.	35	13 : 2	6·87	1·18	0·20	0·14
B.	46	13 : 0	6·78	1·18	0·18	0·12
C.	106	13 : 6	6·04	1·44	0·14	0·10

TABLE IVB

Differences in Mean Attitude Scores and S.D.'s for Groups A, B and C

Group	Mean Score	Diff. in Score	σ_{m_d}	't'	S.D.	Diff. in S.D.	σ_{σ_d}	't'
A............	6·87	0·09	0·27	0·33	1·18	0·00	0·19	—
B............	6·78				1·18			
A............	6·87	0·83	0·25	3·32 *	1·18	0·26	0·17	1·53
C.	6·04				1·44			
B............	6·78	0·74	0·23	3·22 *	1·18	0·26	0·16	1·63
C.	6·04				1·44			

* Significant at the 1 per cent. level.

TABLE VA

MEAN ATTITUDE SCORES FOR GROUPS IN RELATION TO LINGUISTIC BACKGROUND AS MEASURED BY LANGUAGE QUESTIONNAIRE

Group	Definition of Group	No.	Average Age	Mean Score	S.D.	σ_m	σ_σ
I.	Monoglot English with slight Welsh	72	12 : 11	6·89	0·91	0·11	0·08
II.	Monoglot English	115	13 : 7	6·07	1·51	0·14	0·10

TABLE VB

DIFFERENCES IN MEAN ATTITUDE SCORES AND S.D.'S FOR GROUPS I AND II IN TABLE VA ABOVE

Group	Mean Score	Diff. in Score	σ_{m_d}	't'	S.D.	Diff. in S.D.	σ_{σ_d}	't'
I. Monoglot English with slight Welsh	6·89	0·82	0·18	4·56 *	0·91	0·60	0·13	4·62 *
II. Monoglot English ...	6·07				1·51			

* Significant at the 1 per cent. level.

NOTES :

1. σ_m = Standard error of the mean, σ_σ = standard error of the standard deviation, σ_{m_d} = standard error of the mean difference, and σ_{σ_d} = standard error of the difference between two S.D.'s. All standard errors have been calculated by substituting N - 1 for N in the denominator. Cf. LINDQUIST, E. F. : *Statistical Analysis in Educational Research* (Boston : Houghton Mifflin, 1940, pp. 48-50).

2. 't' = Critical Ratio = Difference in Score divided by the standard error of the difference. Its significance has been evaluated by reference to Table III of Fisher and Yates : *Statistical Tables* (Edinburgh : Oliver and Boyd, 1948), entering the table with $(N_1 - 1) + (N_2 - 1)$ degrees of freedom. Cf. also Table C on p. 330 of EDWARDS, ALLEN L. : *Statistical Analysis for Students in Psychology and Education* (New York : Rinehart and Company, Inc., 1946).

V.—DISCUSSION OF RESULTS.

The main points which emerge from the above statistical analysis are as follows :

(1) The mean attitude score for the whole group is 6·37.

(2) Table IA indicates that attitude towards Welsh is most favourable in the youngest age-group, and thereafter steadily declines with advancing age. Table IB shows a statistically significant difference between the first and the fourth year, between the second and the fourth year, and also between the third and the fourth year, whereas no such difference is revealed between the first and the second and third years, nor between the second and the third year. It appears, therefore, that pupils' attitude to Welsh remains more or less constant for the first three years, but that the fourth year is decidedly critical in this respect. Other investigators have discovered a similar tendency in the case of other languages, notably David Jordan[1] in the case of French, and R. A.

---

[1] JORDAN, DAVID : "The Attitude of Central School Pupils to Certain School Subjects, and the Correlation between Attitude and Attainment," this *Journal*, February, 1941, Vol. XI, pp. 28-44.

Pritchard[1] in the case of Latin. Thus R. A. Pritchard notes that a certain amount of interest in Latin mentioned in about 50 per cent. cases in the early age-divisions seemed " to disappear as a reason *above* 14½."[2] It is a significant fact that the age quoted by him exactly corresponds to the age-group in which the most significant decline in attitude to Welsh was observed in the present enquiry, that is, 14 : 6 to 15 : 5. About 50 per cent. cases in this age-division were definitely of the opinion that they did not like Welsh ; their answers to Section E of the test-paper are illuminating, and partly account for the unpopularity of the subject at this stage. Two reasons predominate, namely, the difficulty of the subject, and a feeling that it has no utilitarian value. More than half of the pupils in the fourth year give these two reasons for not liking Welsh.

The difficulty of the subject is sometimes referred to in a general way, for example, " I think it is a difficult language," or " to me it is a hard language to learn and to speak." In some cases, however, specific reference is made to difficulties of grammar and pronunciation. The difficulty experienced by these fourth-year pupils is probably attributable not to a single condition, but rather to a multiplicity of interacting causes, mental, physical and environmental. Further investigations are required to reveal the exact influence on the pupil's proficiency in a second language of such factors as inferior general and special intellectual abilities, adverse temperamental and emotional states, and unfavourable educational, physical and social conditions. Amongst social conditions the verbal and general cultural background of the home and parental attitudes and interests will undoubtedly be factors of primary importance.

The difficulty of the subject also appears much more frequently as a reason for its unpopularity in the first and fourth years than in the second and third years. Such a tendency is probably to be expected in the fourth year, but it is not so easy to explain in the case of the first year. It is tentatively suggested that the difficulty so frequently experienced by pupils in the first-year arises from a lack of continuity in the Welsh course as between the primary and the secondary schools. It is often found that the first-year work in the secondary school is not adequately related to previous knowledge gained at the primary stage.

The other prominent reason noted in the replies was the feeling that Welsh was of no value after leaving school ; it appears much oftener in the fourth year than in the earlier ones, and it is stated with clarity and vigour as the following examples will serve to show : " I don't think it will come in very useful when I leave school," " I do not think it is really necessary," " I will not need it when I leave school." The problem is partly one of bias in teaching-method : the teaching of Welsh in the fourth year in a secondary modern school obviously needs to be planned with an adequate emphasis on its value for after-school life, and every effort should be made to bring the pupils into close and frequent contact with the present-day life of Welsh-speaking communities in other parts of Wales. To teach the language *in vacuo* serves no useful purpose. This feeling of futility and its consequent frustration could also be more successfully combated if Welsh were taught more as a *living* language than as a collection of grammatical rules and precepts ; thus one pupil, bemoaning the fact that Welsh " gets a bit boresome at times," also adds the following significant comment : " I think I would like Welsh better if we could hold a conversation between one another in Welsh." The paramount importance of an oral approach to the subject even in the fourth year is further corroborated by the fact that the number of pupils in the present enquiry who wished to learn enough Welsh to speak it and to understand others speaking it was 10 per cent. higher in the fourth year than in the first year.

---

[1] PRITCHARD, R. A. : " The Relative Popularity of Secondary School Subjects at Various Ages," this *Journal*, June, 1935, Vol. V, Part III, pp. 229-241.

[2] Italics are mine.

Table IB also shows a statistically significant difference in variability between the first and the second year, and the first and the fourth year, and also one which is significant at the 5 per cent. level between the first and the third year. Differences in variability between the second, third and fourth years are not statistically significant. There is, therefore, a significant increase in variability of attitude at the beginning of the second year without any significant increase thereafter.

(3) Tables IIA and IIB indicate that the differences observed between A and B forms as regards mean score and variability of attitude to Welsh are not statistically significant. This finding will be discussed again in a later section of this paper.

(4) There is a sex difference in favour of the girls which is significant at the 1 per cent. level, as shown by the results in Tables IIIA and IIIB. One may conclude that girls show a definitely more favourable attitude to Welsh than boys as tested by these Attitude Tests. A similar superiority of girls over boys was observed in the previous investigation. This finding is in agreement with the results of other researches which report that girls are usually found superior to boys in language studies.[1] It should be noted, however, that observed sex differences of this kind are often treated with the utmost caution by some writers on the subject ; for example, Sandiford emphasies that " these differences must be regarded only as trends ; any study of the summarized findings . . . shows many discrepancies and reversals."[2] Fleming, in this country, affirms that " absence of convincing evidence of significant differences between the sexes characterises. . . . their performance in objective measurements of school status."[3] She is extremely sceptical of many of the reports which claim a higher average for girls in tests involving linguistic skill, because the results have often been obtained from selected samples influenced by differences in the school history of boys and girls. On the other hand, it is equally possible that the true magnitude of sex differences in various scholastic subjects is obscured in some studies by a lack of completeness of ascertainment in the sample tested—even large numbers sometimes prove to be less representative of the *total* population than relatively small numbers which form a complete cross-section of the population within a limited geographical area ; in the latter event it is also much easier to ensure that schooling conditions are as nearly identical as it is possible to arrange them for the two sexes. There appears, therefore, a great deal of conflicting and inconsistent evidence as regards sex differences in educational achievement ; *for practical purposes*, however, there is no doubt that individual differences *within one sex* are far more important than the differences *between the sexes* which have hitherto been observed;[4] and the sex difference in respect of attitude to Welsh in the present results should be regarded accordingly by practising teachers. No statistically significant difference was observed between the sexes in respect of variability of attitude.

---

[1] Cf. THORNDIKE, E. L. : *Educational Psychology, Vol. III* (Teachers' College, Columbia University, N.Y., 1925, p. 184). SANDFORD, PETER : *Foundations of Educational Psychology* (Longmans, Green and Co., 1938, p. 150). Also MACGREGOR, GREGOR : *Achievement Tests in the Primary School* (Publications of the Scottish Council for Research in Education, VI, University of London Press, 1936, pp. 75-86,) where a significant difference in favour of the girls is reported in the case of language usage and spelling, and one in favour of the boys in reading and arithmetic reasoning.

[2] *Op. cit.*, p. 150.

[3] FLEMING, C. M. : *Adolescence* (London : Routledge and Kegan Paul, Ltd., p. 100).

[4] Cf. THORNDIKE, *op. cit.*, p. 184.

(5) The linguistic background of the home provides a potent source of influence on attitude to Welsh. This fact is amply illustrated by Tables IVA and IVB. On the basis of the supplementary question attached to the Language Questionnaire pupils were divided into three sub-groups as follows: Group A=those whose father or mother, or both father and mother, spoke Welsh; Group B=pupils whose father or mother, or both father and mother spoke a little Welsh, and Group C=pupils of non-Welsh-speaking parents. A significant difference in mean attitude score was found between Groups A and C, and also between Groups B and C. No significant difference was observed between Groups B and C. Differences in variability of attitude between these three groups were not statistically significant. The influence of language conditions outside school on attitude to Welsh is even more significant when the pupils are divided into two sub-groups on the basis of their replies to the Language Questionnaire. Group I consisted of those pupils who may be described as monoglot English with slight Welsh, that is, pupils whose score on the 24-unit scale of marking ranged from 23 for English and 1 for Welsh to 19 for English and 5 for Welsh. Group II was composed of monoglot English pupils, whose score was 24 for English and 0 for Welsh. The difference in mean attitude score in favour of Group I was highly significant at the 1 per cent. level. The difference in variability of attitude observed between these two groups was also highly significant at the 1 per cent. level.

(6) Responses to Section D of the test-paper were carefully summarised and a statistical analysis prepared, indicating the number of pupils per cent. with reference to their preferences for various aspects of the study of Welsh. 15·71 per cent. did not want to learn any Welsh at all, 21·43 per cent. boys, and 10·71 per cent. girls. The number per cent. of pupils in this category increased steadily from year to year, 1st year=6·67, 2nd year=13·56, 3rd year=16·98, and 4th year=31·58. The increase from year to year was also more pronounced for boys than for girls. 6·19 per cent. of the total number of pupils only wanted to learn enough Welsh to understand others speaking it, 9·18 per cent. boys, and 3·57 per cent. girls. 22·38 per cent. of the entire group wanted to be able to speak Welsh themselves as well as to understand others speaking it; there was a definite increase from year to year in the number of pupils in this category, with a slight decrease in the fourth year—the figures were as follows: 1st year= 16·67 per cent., 2nd year=20·34 per cent., 3rd year=28·30 per cent., and 4th year= 26·32 per cent. A preponderance of boys over girls was also observed in the first two years. 1·91 per cent. only wanted to learn enough Welsh to be able to read it, and 3·33 per cent. wished to learn enough Welsh to be able to write it as well. 50·48 per cent. of the whole group wanted to learn Welsh in all its aspects, the active and passive uses of both the oral and written forms; the figures for boys and girls, respectively, were 40·82 per cent. and 58·93 per cent. There was a marked decline in the number per cent. of pupils in this category from year to year, as shown by the following figures: 1st year=66·67, 2nd year=55·93, 3rd year=41·51, and 4th year=28·95. The rate of the decline was also much more evident in the case of boys.

(7) The statements of reasons for or against learning Welsh received from the pupils in answer to Section E of the test-paper supplied many interesting and suggestive data. It is not possible to discuss them fully in this paper. 164 pupils liked Welsh, and 47 pupils did not like the subject. Only a small number, 14·22 per cent., the majority of whom were boys, failed to give an adequate reason for their choice; the reasons offered in these instances were mostly general in character—for example, " I like Welsh because it is interesting," " I do not like Welsh because it is so boring in my opinion." It was also

observed that unsatisfactory reasons of this kind were confined to pupils in the first three years.

Answers favourable to learning Welsh were classified according to the character of the reason given. Four types of reasons predominated. The most important factor deciding pupils' choice was "interest"; the second most important factor was the "utilitarian" value of the subject; reasons of a "national" or "racial" character came third in the list of reasons, and "proficiency" in the subject was a poor fourth. E. O. Lewis[1] likewise found that utilitarian, interest, and proficiency reasons, in that order of importance, were major factors in deciding pupils' choice of subject. Liking for the teacher, family attitudes to the subject, and the fun experienced in the lessons were sometimes mentioned as reasons. There were also a few reasons which could not be easily brought within any system of classification; some of these were original and spontaneous in character, for example, "I like Welsh because it is a change from English." It was also found that the four predominant reasons already mentioned varied slightly in order of importance from year to year. Thus interest in the subject was easily the most important factor in the first three years, whilst in the fourth year, together with proficiency, it made a poor third in the list. Reasons of a utilitarian character were second in importance throughout the first three years, but obtained a clear majority in the fourth year—the explanation of this preponderance of the utilitarian reason in the fourth year is probably to be found in the fact that these pupils were about to leave school, and were, therefore, more prone to evaluate a subject with reference to the choice of a vocation in the near future. It is evident, moreover, that this utilitarian factor has considerable influence at all stages of the course; this fact has great pedagogical significance, and undoubtedly suggests that teachers should utilise this practical outlook of the child to the fullest possible extent.

The usefulness of Welsh is sometimes referred to in general terms, for example, "It is quite a useful language." In most cases, however, the term is defined in more specific detail, for example, "It is useful because people speak it in Wales," "It will be very handy in a Welsh town," "When it is time to leave school I might go to a Welsh Eisteddfod." Almost invariably the conception of Welsh as a means of communication with Welsh-speaking people in other parts of Wales enters into the pupils' utilitarian reasons.

Reasons of a national or racial character were fourth in order of importance in the first year, third in the second and third years, and second in the fourth year, thus increasing in importance from year to year. One or two examples will suffice to illustrate the nature of this type of reason: "I like Welsh because it is the language of my country," "I like Welsh because when you listen to it, it gives you the true feeling of a Welshman," "I like Welsh because I belong to Wales."

Proficiency as a reason for liking Welsh was third in order of importance in the first year, and equally third with interest in the fourth year; in the second and third years it was rather low in the list.

The reasons given by those who do not like Welsh are equally instructive. The difficulty of the subject, its lack of utilitarian value, and insufficient interest in it are almost equally prominent. The difficulty of the subject is the most prominent reason in the first year, and insufficient interest in the subject in the third year. A more or less equal balance between these three major reasons is maintained in the second year. In the fourth year the subject's lack of utilitarian value easily predominates over all other reasons. Some of the pupils express a desire to devote the time given to Welsh to some other subject, such as French, or any other foreign language, and one pupil writes : " I should like to have more practical lessons instead of it. Something of use to me after school." Another factor often mentioned in their answers is the disadvantage of having no Welsh at home, for example, " Neither my father nor my mother speak Welsh, so I get no help from them."

---

[1] LEWIS, E. O.: "Popular and Unpopular School Subjects," *The Journal of Experimental Pedagogy and Training College Record*, Vol. 2, No. 2, p. 97.

(8) An analysis of the answers given by the pupils to each one of the attitude statements reveals some interesting results. Only the more significant ones are mentioned here. 10·90 per cent. of the 211 pupils would like to stop Welsh from being taught in schools—in the fourth year 28·95 per cent. were of this opinion. 52·61 per cent. would like to spend a little more time on Welsh—the number per cent. declining steadily and abruptly so in the fourth year as follows : 1st year=63·37, 2nd year=55·93, 3rd year=50·00 and 4th year=34·21 per cent. 17·54 per cent. thought that far too much time is given to Welsh in school—this percentage was fairly constant for the first three years, but increased to 28·95 per cent. in the fourth year. 69·19 per cent. thought that everyone in a Welsh school should be taught Welsh, though only 44·74 per cent. were of this opinion in the fourth year as compared with 80·00 per cent. in the first year. 14·22 per cent. thought that Welsh was the least interesting subject that they had in school, and 34·12 per cent. regarded it as the most interesting subject. 94·31 per cent. found some parts of their work in Welsh interesting, and that percentage was more or less maintained throughout the four years. 31·75 per cent. stated that when they left school, they would, give up the study of Welsh entirely—the number per cent. in this respect increased considerably from year to year : 1st year=16·67, 2nd year=23·73, 3rd year=42·59, and 4th year=52·63. 35·07 per cent. considered that Welsh grammar was uninteresting—this dislike of grammar was most pronounced in the last year, and was expressed by 57·89 per cent. of these fourth-year pupils. On the other hand, conversational lessons in Welsh were favoured by 73·93 per cent., the percentages for the first three years, respectively, being 75·00, 74·38 and 79·63 ; even in the fourth year the percentage was as high as 63·16. 63·37 per cent. of the pupils in the first year tried to do something every day to improve their Welsh vocabulary ; in the second year 57·63 per cent. did so, in the third year 33·33 per cent., and only 15·79 per cent. in the 4th year. 30·00 per cent. in the first year, 28·81 in the second year, 12·96 in the 3rd year, and 2·63 in the fourth year stated that if they were given the choice of a book, they would choose a Welsh one. These figures undoubtedly corroborate the tendency already noted for attitude to Welsh to decline with increasing age ; they also emphasise the importance of an oral rather than a grammatical and formal approach to the subject, *even in the fourth year of the course,* and indicate the urgency of arresting the decline in vocabulary from year to year, and of providing suitable and well-graded reading material for pupils in the secondary modern school.

Finally, it was found that 61·61 per cent. of the pupils expressed a desire to spend some time in a Welsh-speaking part of Wales in order to get more practice in speaking Welsh. It is suggested that teachers might further explore this possibility with considerable advantage to the teaching of Welsh as a second language in the English-speaking parts of Wales.

VI.—The relationship between attitude and intelligence.

It will be remembered that no significant difference in respect of attitude to Welsh was found between A and B forms in the present enquiry ; a similar result was obtained in the preliminary investigation in the Rhondda Valley. It was possible on this occasion to make a further study of this finding, since all pupils on entry into the school are tested for intelligence by the headmaster—the results so obtained are used for classifying the pupils into three streams, A, B and C. Recent intelligence-test results were available in respect of the first-year pupils tested for attitude, and the headmaster kindly permitted the writer to use them in connection with the present investigation. The group intelligence test used was the Advanced Otis—Form A. An examination of the results showed that the performance of a child had been represented by the Index of Brightness or I.B., this figure being 100 ± the deviation from the American norm. The same term has been used in the present enquiry, but the I.B.'s were recalculated on the basis of more accurate British norms,[1] and also with an additional adjustment for the association between variance and age. A detailed account of the procedure for adjusting for variation in variance and the figures necessary for the calculation are given in an article by Fraser Roberts and others.[2]

The following tables summarise the results :

MEAN INTELLIGENCE SCORES FOR A AND B FORMS, FIRST YEAR

Forms	No.	Average Age	Mean I.B.	S.D.	σ_m	σ_σ
IA	32	12 : 9	106·03	13·81	2·48	1·75
IB	26	12 : 1	87·46	16·09	3·22	2·28

DIFFERENCES IN MEAN INTELLIGENCE SCORES AND S.D.'S FOR A AND B FORMS, FIRST YEAR

Forms	Mean I.B.	Diff. in Mean	σ_{m_d}	't'	S.D.	Diff. in S.D.	σ_{σ_d}	't'
IA	106·03	18·57	4·06	4·57*	13·81	2·28	2·87	0·79
IB	87·46				16·09			

* Significant at the 1 per cent. level.

A statistically significant difference is observed between the two groups in respect of mean intelligence score, whereas no such difference was found between them in respect of attitude to Welsh. The difference in intelligence between the two groups does not seem to have a comparable influence on their attitude. This result is further corroborated by the correlation observed between attitude and intelligence in the combined results of the two groups. The coefficient, obtained by the Pearson-Bravais method, was −0·119, which is insignificant even at the 5 per cent. level.[3] Drummond, who used attitude tests of a verbal and pictorial kind to investigate the attitudes of backward pupils towards school life and work, found positive correlations between attitude and intelligence.[1] On comparing normal, backward and dull groups by examining the differences between mean attitudes by means of an analysis of variance, she also found certain significant differences between normal and dull pupils and between backward and dull pupils.[2] It is obviously necessary to repeat investigations of this kind with larger numbers—unfortunately, statistical data were not available for measuring the relationship between attitude and intelligence in the case of the remaining groups in the present enquiry.

---

[1] FRASER, ROBERTS J. A., NORMAN, R. M., and GRIFFITHS, R. : "Studies on a Child Population," *Annals of Eugenics*, Vol. VI, Part IV, December, 1935, p. 332.

[2] *ib.* pp. 333–334.

[3] The significance of this and subsequent r's has been evaluated by reference to Table VI of FISHER and YATES : *Op cit.*, the number of degrees of freedom available being N - 2, where N is the number of pairs. Cf. also Table D on p. 331 of EDWARDS, ALLEN L. : *Op. cit.*

VII.—THE CORRELATION BETWEEN ATTITUDE AND ATTAINMENT.

Results obtained by other investigators indicate a positive relationship between attitude and attainment.[3] In the present enquiry the calculation of the correlations between these two factors was made much easier by the fact that the attitude test was given shortly before the terminal examination; the teaching of all pupils by the same specialist teacher was also a distinct advantage in this respect. Since the terminal examination papers were set to meet the requirements of each year separately, the correlations were calculated from year to year, and more adequate numbers were thus made available. The coefficients, calculated by the Pearson-Bravais method, were as follows:

THE MEASUREMENT OF CORRELATION BETWEEN ATTITUDE AND ATTAINMENT IN WELSH

Year	I	II	III	IV	Mean
No.	53	46	48	33	180
r.	+0·24	+0·12	+0·56*	+0·39†	+0·33

* Significant at the 1 per cent. level.
† Significant at the 5 per cent. level.

The results suggest a slight but positive relationship between attitude and attainment in Welsh in the first two years, whereas in the third and fourth years this relationship has assumed significant proportions. It appears that Jordan's

---

[1] DRUMMOND, G. A.: "The Attitude of Backward Adolescents towards their School Life and Work" (M.A. Thesis, Univ. of London, 1947), and quoted by SCHONELL, F. J.: "The Development of Educational Research in Great Britain, Part III," this *Journal*, Vol. XVIII, Part II, June, 1948, p. 56. The following correlations are given:

	Attitude (lessons)	Attitude towards School Work
Mental Age	·33 ± ·065	·37 ± ·063
I.Q.	·40 ± ·062	·35 ±065

[2] SCHONELL, F. J.: *Ib.*, p. 56.

[3] Cf. JORDAN: *Op. cit.*, p. 41. Using Spearman's Rank Order Method he obtained the following correlations between attitude and attainment in French, as represented by a composite mark based upon the child's examination result and his work during the term:

Forms	IR	IA	IB	2A	2B	3A	3C	3N	Mean
ρ	—·04	+·18	+·15	+·29	+·34	+·47	+·40	—	+·26

Drummond, quoted by Schonell (*op. cit.* p. 56), found the following correlations between educational attainment and attitudes.:

	Attitude (lessons).	Attitude to School Work
Educational Age	·25 ± ·069	·29 ± ·067
Educational Quotient	·27 ± ·069	·34 ± ·066

correlations likewise tend to increase from year to year.[1] It is emphasised, however, that the present results are only tentative in character, pending further investigations with specially constructed and adequately standardised tests for measuring attainment in Welsh as a second language.[2]

VIII.—SUMMARY AND CONCLUSIONS.

The purpose of the present investigation of attitude towards Welsh as a second language was to obtain further evidence either for or against the differences and tendencies previously observed in a preliminary enquiry undertaken in a secondary modern school in the Rhondda Valley. Improved scales for measuring attitude and linguistic background were given to 211 pupils of 11 : 6 to 15 : 5 years of age grouped in eight different forms in a secondary modern school in a suburban area outside Cardiff. The test-sheets, in addition to attitude statements, also contained material whereby pupils' reasons for or against learning Welsh and their preferences in relation to various aspects of the process could be ascertained. All pupils were taught Welsh on the Direct Method by the same specialist teacher of Welsh.

The present enquiry evidently corroborates the findings of the preliminary investigation in the following respects :

1.—Attitude to Welsh becomes less favourable with increasing age. The present results, however, indicate that the differences observed between the first three years are insignificant, whereas in the fourth year the decline in attitude has assumed statistically significant proportions.

2.—A and B forms do not differ significantly from one another either in average attitude score or in variability.

3.—There is a statistically significant sex difference in the results. Girls show a more favourable attitude than boys. There is conflicting evidence in respect of variability of attitude between the sexes.

4.—The linguistic structure of the out-of-school environment is a potent source of influence on attitude to Welsh. Monoglot English-speaking pupils with a slight knowledge or experience of Welsh outside school show a more favourable and less variable attitude than monoglot English-speaking pupils whose acquaintance with Welsh is confined to the classroom.

5.—Formal grammar lessons appear to be very unpopular at all stages of the course, and especially so in the last year. Conversational lessons, on the other hand, are favoured by a large majority of the children, and interest in this oral approach to the subject is maintained at a high level throughout the course.

6.—There is an indication of a steady decline of interest in Welsh reading and vocabulary from year to year.

Further results obtained in the present enquiry may be briefly summarised as follows :

1.—Only 15·71 per cent. of the pupils had no desire at all to learn Welsh, whereas 84·29 per cent. declared themselves to be in favour of learning the language. Exactly one half of the entire group wished to learn Welsh in all its aspects, that is, the active and passive uses of both its oral and written forms, about one-third was interested in the oral aspect only, and only about one-twentieth in the active and passive uses of the written form of the language.

---

[1] See footnote 3 on page 130.

[2] Educational research of a statistical nature is inevitably held up in Wales by the almost complete absence of *standardised* mental tests suitable for use in Welsh-speaking areas.

2.—Four predominant reasons were offered on behalf of learning Welsh; "interest" was the first on the list of reasons, the "utilitarian" value of the subject second, reasons of a "national" or "racial" character were third, and "proficiency" in the subject was fourth. This order varied slightly from year to year—thus "interest" was the most prominent factor in the first three years, but in the fourth year the "utilitarian" reason obtained a clear majority.

The difficulty of the subject, its lack of utilitarian value, and insufficient interest in it are almost equally prominent in the list of reasons against learning Welsh.

3.—The statistically significant difference in mean intelligence scores between A and B forms in the first year was in no way reflected in their respective attitude scores. Furthermore, a negative but statistically insignificant correlation was found between attitude and intelligence in the combined results of these two forms.

4.—In the first two years slight but positive correlations were found between attitude and attainment in Welsh as represented by the pupils' performance in a terminal test paper; in the third and fourth years the correlations between these two factors were statistically significant. This finding seems to suggest that attainment in Welsh is increasingly influenced by the pupil's attitude in the later stages of the course.

The writer desires to thank Professor C. W. Valentine for very helpful suggestions, Professor Sir Ifor Williams and Mr. E. B. Simpson for reading the paper in typescript, and offering valuable advice, Mr. G. Eifion Williams for assistance in checking calculations, and the headmaster and specialist teacher of Welsh at the school for kindly co-operating in the investigation.

THE POSITION AND SIZE OF THE BORDER-LINE GROUP IN AN EXAMINATION

By A. E. G. PILLINER
(Moray House, University of Edinburgh).

I.—*The problem.* II.—*The statistical method.* III.—*The position and size of the border-line group.* IV.—*Method of constructing the table.* V.—*Summary.* VI.—*References.*

I.—THE PROBLEM.

IN assessing the suitability of candidates for grammar school education in the 11+ examination, a number of education authorities examine further a group of candidates at or near a border-line decided on the basis of previous experience. An attempt is made in this article to estimate the position and size of the border-line group on statistical grounds.

Most authorities use a battery comprising several tests, for example, one Intelligence, one English and one Arithmetic test. The scores on these are standardised and the sum of the resulting "quotients" computed for each candidate.

Were all forms of error absent, that is, were the battery perfectly reliable, the problem would resolve itself into choosing the candidates whose quotients exceeded the border-line. As no battery is perfectly reliable, a degree of uncertainty attaches to the quotients obtained by candidates, and it may happen that in the region of the border-line a candidate whose obtained quotient falls on one side of the line has a true quotient on the other side. The degree of uncertainty increases as the reliability decreases.

II.—THE STATISTICAL METHOD.

The battery reliability is the correlation that would be expected between two applications of two parallel batteries to the same group of children. It is based on the test-retest correlations and inter-correlations of the various tests comprising the battery and can be obtained from these by the use of a pooling square. For a battery of Moray House Intelligence, English and Arithmetic tests, equal weight being attached to each test, the reliability thus calculated is ·983.

If the standard deviations of the tests in the battery are known (and in this case they are, since all have been made equal to 15 by standardisation), the pooling square gives the standard deviation of the summed quotients for the battery.

From these statistics the standard error of the quotient sums for individual children may be calculated. Knowing in addition the sum of the mean quotients for the authority and the percentage of the year-group that is to be admitted to the grammar school, it is possible to estimate (i) the position of the border-line to secure this admission: (ii) the limits of the border-line group and (iii) the size of the border-line group for further examination, expressed as a percentage of the year-group. The method is given in Section IV.

III.—THE POSITION AND SIZE OF THE BORDER-LINE GROUP.

The table below gives (i), (ii) and (iii) above for those authorities using one Intelligence, one English and one Arithmetic Test, each receiving the same weight. The national mean is 100 for each test, the standard deviation is 15, and the obtained scores are converted into quotients using national norms as published by the University of London Press. The grammar school percentage intake is shown horizontally along the top of the Table, and the sum of mean quotients vertically on the left.

Sum of Mean Quotients within Authorities	Percentage to be Admitted	10%	15%	20%	25%	30%
	Border-line Group	9%	12%	14%	16%	18%
275	A	330	319	311	304	297
	B	319—341	308—330	300—322	293—315	286—308
280	A	335	324	316	309	302
	B	324—346	313—335	305—327	298—320	291—313
285	A	340	329	321	314	307
	B	329—351	318—340	310—332	303—325	296—318
290	A	345	334	326	319	312
	B	334—356	323—345	315—337	308—330	301—323
295	A	350	339	331	324	317
	B	339—361	328—350	320—342	313—335	306—328
300	A	355	344	336	329	322
	B	344—366	333—355	325—347	318—340	311—333
305	A	360	349	341	334	327
	B	349—371	338—360	330—352	323—345	316—338
310	A	365	354	346	339	332
	B	354—376	343—365	335—357	328—350	321—343
315	A	370	359	351	344	337
	B	359—381	348—370	340—362	333—355	326—348
320	A	375	364	356	349	342
	B	364—386	353—375	345—367	338—360	331—353
325	A	380	369	361	354	347
	B	369—391	358—380	350—372	343—365	336—358

A=Border-line Score.
B=Limits of Border-line Group.
Values not shown in the Table can be obtained by interpolation.
If tests are standardised on obtained scores, so that the mean for each is 100, use row 300.

It should not be thought that the inclusion of entries in the top right hand corner of the body of the table implies approval of children of this calibre proceeding to the grammar school. The figure 297, for example, implies a mean quotient of 99 for a child. However, this part of the Table may prove useful for other types of secondary school.

IV.—METHOD OF CONSTRUCTING THE TABLE.

(1) From data compiled over a period of years from Moray House Tests, the following correlations are obtained :

Variables Correlated	Intelligence -Intelligence	English -English	Arithmetic -Arithmetic	Intelligence -English	Intelligence -Arithmetic	English -Arithmetic
Correlation Coefficients	·940	·961	·961	·897	·859	·815
Number of Cases ..	5409	4401	3611	3178	3423	2643
Inverval between Tests	41-280 days	41-84 days	41-48 days	0-2 days	0-2 days	0-2 days

Each correlation is the result of pooling a number of observed coefficients, using Fisher's Z technique (2, p. 204).

(2) The reliability of the battery I+E+A is found, using a pooling square, to be ·983 (3, p. 85).

(3) The pooling square gives also the standard deviation of the battery I+E+A as 42·8 (4, p. 11).

(4) The standard error of a child's score on the battery is :

$$42{\cdot}8\sqrt{1-{\cdot}983}=5{\cdot}6 \quad (1,\ \text{p. } 407).$$

(5) The confidence limits for the border-line at the 5 per cent. probability level are :

border-line score $\pm 1{\cdot}96 \times 5{\cdot}6$ (5, table 15).

i.e., approximately 11 points on either side of the border-line.

Thus all candidates with quotient sums lying within this range should be presented for further examination.

The 5 per cent. level was chosen after some consideration as a compromise. The limits given will include 19 out of every 20 children who might merit re-examination. To include more than these would result in a border-line group which it would be scarcely practicable to re-examine.

(6) The position of the border-line is determined by the percentage intake of the grammar schools. If this is to be 10 per cent., for example, the line will lie at 1·2816 standard deviations, or 54·85 points, above the mean. If the latter is, say, 275, the border line is at 329·85, or 330 to the nearest whole number (6, Table 1).

(7) Since the variate is summed quotients on standardised tests, it is normally distributed about the authoritys' mean summed quotient (based on national norms) with a standard deviation of 42·8. The percentage of the whole year-group lying in the border-line group is calculated from Pearson's Table II (5, Table 15).

(8) It will be observed that the border-line group increases in size as the grammar school intake increases. This is because the border-line is being pushed towards the mean of the normal curve. The area under the curve between the limits is, therefore, increased.

V.—SUMMARY.

The fact that no test or battery of tests is perfectly reliable introduces a degree of uncertainty into the score obtained on the test or battery of tests. This should be taken into account by education authorities in their examination for entry into grammar schools. A table is constructed to assist in assessing the position and the size of the resulting border-line group for various sizes of grammar school intake. The method used in constructing the table is outlined briefly.

VI.—REFERENCES.

1. T. L. KELLEY : *Fundamentals of Statistics* (Harvard, 1947).
2. R. A. FISHER : *Statistical Methods for Research Workers* (Oliver and Boyd, 1946).
3. SIR GODFREY THOMSON : *The Factorial Analysis of Human Ability* (U.L.P., 1946).
4. G. A. FERGUSON : *The Reliability of Mental Tests* (U.L.P., 1947).
5. H. E. GARRETT : *Statistics in Psychology and Education* (Longmans, 1937).
6. KARL PEARSON : *Tables for Statisticians and Biometricians*, Part I (C.U.P., 1930).

The writer wishes to thank Mr. W. G. Emmett, M.A., for helpful advice and criticism in the preparation of this article.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES.[1]

A Study in Classroom Cheating in Arithmetic and Spelling.[2]

BY JOSEPH McNALLY.

(*Summary of thesis submitted in part fulfilment of the requirements for the Degree of Bachelor of Education at Glasgow University*, 1949.)

A series of experiments was carried out on children's cheating, making use of an arithmetic and a spelling test. The main purpose of the work was to obtain some data bearing on the problem whether pupils should be allowed to correct their own work, and the measurement and consistency of honesty were points of central interest.

Method.—The tests chosen for cheating detection were: the Moray House Arithmetic test, and Burt's Spelling test on which the 50 words covering age groups 10-14 were given as one unified list.

Technique.—May and Hartshorne's "duplicating technique" was used. The tests were administered on two occasions, on the first occasion as ordinary arithmetic and spelling tests, and on the second occasion—after one or two days—the same tests were returned, the pupils being asked to score their own papers.

Scoring.—The difference between the real achievement, on the first administration, and the self-marked paper was called the "amount" score of cheating. The observation that a cheating response was made was called the "fact" score.

Situations.—Cheating was studied in two distinct situations: (1) Social, in which the pupils were asked to shout out in the classroom their total mark for arithmetic and spelling; (2) Private, in which the pupils were asked to keep their mark to themselves.

Subjects.—The test was administered to all first-year classes in a city senior secondary school. The pupils of both sexes came mainly from working and lower middle classes.

The tests were spread over two years. In the first, the research was of a preliminary nature, in order to test the technique and gain experience. In the second, the tests were applied to 312 pupils (144 boys and 168 girls), in both situations, social and private.

RESULTS.

First Year.—Preliminary Experiment (213 subjects):

(1) Cheating is very prevalent: 82 per cent. in Arithmetic and 87 to 91 per cent. in Spelling.

(2) Cheating in Spelling is greater than in Arithmetic, the difference being 2·93 times its S.E.

Second Year.—Main Experiment (312 subjects):

The tests were administered twice under same conditions. The grand mean of the "amount" score (the mean of the differences on the first and second occasion) in Arithmetic was 3·11 (S.D. 4·11) and in Spelling 5·24 (S.D. 4·73). The following conclusions were drawn:

(1) The difference between the test results on the first and second occasions was not significant.

(2) The cheating was greater in Spelling than in Arithmetic, because there was an "amount" score of 5·24 against 3·11, or, in standard scores, 1·11 for Spelling and 0·76 for Arithmetic.

---

[1] These Outlines must be submitted through the Head of the Department in which the research was carried out.

[2] Summarised by Z. Barbu.

Analysis of Results of Main Experiments.—The following formulæ were used in the analysis of variance :

(1) $\Sigma(x-\bar{x})^2$ for total variance.

(2) $\Sigma[n_c(\bar{x}_c-\bar{x})^2]$ for between classes variance.

(3) $\Sigma\,\Sigma(x-\bar{x}_c)^2$ for within classes variance (between the individuals of the same class).

The between-class variance, $z=0{\cdot}2828$, was significant (over $0{\cdot}2000$) but this difference was due to Class J only, and the difference between this Class and the whole group was 2·8 times its own S.E. All other between-class differences were insignificant. If the Class J is eliminated the difference between classes is insignificant ($z=0{\cdot}1450$), while the difference within classes is significant ($z=3{\cdot}366$). In the Spelling test the between-class difference gave $z=0{\cdot}3244$. If the rebel class (D) is excluded, z dropped to 0·1265, and became insignificant.

Sex differences.—Using the " fact " score, the following data were found :

DEGREE OF HONESTY

	Honest on both tests	Honest on one test	Never Honest	Total
Boys	36	37	71	144
Girls	38	46	84	168
Total	74	83	155	312

No sex difference was found ($\chi^2=0{\cdot}276$). Similarly no difference was found between Private and Social situations ($\chi^2=4{\cdot}346$).

Specificity versus Generality.—There was no significant consistency in responses. For example, from 121 pupils who gave honest responses in the first Arithmetic test, only 74 continued to be honest in the second Arithmetic test, and of these, only 32 remained honest in both Spelling tests. Even when the situation was similar, the pupils did not show consistency in their responses. When the situations differed, the inconsistency increased.

Correlations between Arithmetic test on first and second occasion ; between Spelling test on first and second occasion, and between Arithmetic and Spelling on first occasion, were as follows and all were significant. $r_{A1/A2}=\cdot 624$; $r_{S1/S2}=\cdot 602$ and $r_{A1/S1}=\cdot 446$. The difference between the correlation between the same tests on several administrations and between different tests was ·2166 (S.E. ·08).

Although the between classes difference was not significant, the conception of honesty as a trait specific to the specific situation is valid. From the results of one test, one cannot tell how any one individual will behave in another, even when the test situations are similar.

Effect of Previous Primary School.—Pupils coming from ten primary schools showed no significant difference between schools in the Arithmetic test ($z=0{\cdot}1331$). From a comparison of the results obtained in two consecutive years, the conclusion was drawn that the standard of honesty improved. The mean " amount " score of cheating decreased from 5·15 to 3·11 in Arithmetic, and from 8·12 to 5·24 in Spelling.

Cheating and Achievement.—From the analysis of achievement scores in comparison with cheating scores it appeared that there is a slight tendency for those of low achievement to cheat more than those of high achievement.

Cheating and Sanctions.—Before the administration of the Arithmetic test the following sanctions were applied. :

(1) [illegible] virtue of honesty.

(2) [illegible] that honesty is the best policy.

(3) [illegible] of (*a*) an " isolate " child preferred by no one, and (*b*) a leader or popular pupil.

The results were that only the punishment of a popular pupil, and religious sanction produced, an improvement in honesty.

Cheating and Perseveration.—Cattell's Perseveration test for children was applied to one class group and the results were related to cheating. A coefficient of mean square contingency of ·450 was found, which showed that the perseveration generally occurs among both honest and dishonest groups in fairly equal measure. There is a slight indication that perseveration is higher when cheating is more common.

BIBLIOGRAPHY

Barbu, Z.: *The Function of Lying in the World of the Child* (Cartea Romaneasca, Cluj, Roumania, 1937).

Barbu, Z.: *Contributions to the Psychology of Honesty* (Institutul de Psichologie, Cluj, Roumania, 1939).

Bathurst, J. E.: *et al*, "Written Tests of Honesty" (*Pub. Person. Stud.*, 1927, No. 7, 98-106).

Bird, C.: "The Detection of Cheating in Objective Examinations" (*School and Society*, 1927, No. 25, 261-2).

Bird, C.: "An Improved Method of Detecting Cheating in Objective Examinations" (*J. Educ. Res.*, 1929, IX, 341-348).

Butterfield, E. W.: "Who is Dishonest"? (*School and Society*, 1929, No. 29, 360-362).

Cady, V. M.: "The Estimation of Juvenile Incorrigibility" (*J. Delinquency*, Monograph No. 2, 1923).

Chambers, E. V.: "A Study of Dishonesty among Students of a Parochial Secondary School" (*Ped. Sem.*, 1926, XXXIII, 717-728).

Grundlach, R.: "A Method for the Detection of Cheating in College Examinations" (*School and Society*, 1925, No. 22, 215-16).

Hartshorne, Hugh and May, Mark, A.: *Studies in Deceit* (Macmillan, 1928).

Hartshorne, Hugh and May, Mark, A.: *Studies in the Organisation of Character* (Macmillan, 1930).

Persing, K. M.: "Morals and Chemistry" (*Ed. Rev.*, 1926, 72, 164-8).

Raubenheimer, A. S.: "An Experimental Study of Some Behaviour Traits of the Potentially Delinquent Boy" (*Psych. Monog.*, XXXIV, 6, No. 159, 1925).

Tuttle, H. S.: "Testing the Curriculum in its Natural Setting" (*Religious Education*, 1929, XXIV, 173-5).

Voelker, P. F.: "An Account of Certain Methods of Testing of Moral Reactions in Conduct" (*Religious Education*, 1921, XVI, 81-3).

An Investigation into a System of Moral Instruction.

By DONALD McPHERSON.

(*Summary of thesis submitted in part fulfilment of the Regulations for the Degree of Bachelor of Education, Glasgow University*, 1949.)

An investigation was carried out into the effectiveness of a particular system of Moral Instruction, especially designed—

(*a*) To improve significantly, over a short period of time, the moral judgment of children;
(*b*) To measure objectively the amount of improvement taking place;
(*c*) To correlate the effect of improvement in moral judgment (if any) with the actual conduct of the children concerned.

More practical experimental work in moral instruction is essential. The amount of lip-service paid to moral instruction is high; the amount of practical service—[illegible] amount of help and guidance in solving such problems as finding the best [illegible] how to put them into effect—has been low. To behave consistently in the " right " way, the child must know what the " right " way is ; morality involves both a practical and a theoretical side. The modern negativism which decries the existence of any permanent human values must be combated vigorously because it is this unfounded scepticism which is threatening to rob modern man of his faith and hope.

Formal moral instruction in the past, or at least until very recently, has been subject-centred—abstract, unrealistic, dogmatic, and " sermonizing," with the child playing a passive role. Moral instruction, to be effective, must be child-centred—concrete, meaningful and interesting, with the child playing an active, heuristic role.

The key to moral instruction is the problem situation :

(*a*) The problem situation emphasises the basic factor in moral training and instruction, the resolving of the conflict of values.

(*b*) The setting of the problem situation must be realistic, meaningful and, if possible, appealing, to the child.

(*c*) The design of the problem situation should be of such a nature that the predominant motives, emotions, and values at stake are readily detectable. The designer should be clear in his own mind what motives, emotions, and values he is emphasising.

As the ultimate value of moral instruction is its influence on overt behaviour, techniques should be devised which are capable of correlating moral instruction and overt behaviour in some practical way.

The above points were borne in mind in the design and execution of the experiment under consideration. Four schools were selected (School 1—above average, Schools 2 and 3—average, School 4—below average, for socio-economic level) ; two classes in each, one class the control group and one the experimental group. Both groups were given moral judgment questionnaires, which also included " Confessions " Tests and " Cheating' ' Tests. The experimental group were now given systematic indirect moral instruction—15 minutes every alternate school day for eight instructional periods. The instruction consisted of practice in the solution of problem situations, each corresponding to a similar problem in the original questionnaire, involving the conflict of essentially the same motives, emotions, and values. At the conclusion of the experimental period, the experimental and control groups were given another questionnaire. The two questionnaires were now treated statistically and the main conclusions reached were :

(*a*) The moral judgment of the children of the experimental group was significantly improved as a result of the technique employed.

(*b*) The behaviour of the children, as assessed by Cheating Test and Confessions Test performances, remained however unaffected.

Environment, including Socio-Economic Status.

(*a*) On the basis of differences between Schools 1 and 4, we may conclude that differences in environment, including socio-economic status, produce differences in moral values.

Environmental differences were not critically analysed in the present experiment, but a superficial study of the evidence (i.e., School 1 and 4 differences) suggested, among other things, that among children of below average socio-economic status ideas on punishment and material damage tended on the whole to be less reasoned, more instinctive and unjust, presumably because of example and personal experience.

(*b*) Some values, however, appear to remain constant, irrespective of environment and socio-economic status. These include kindness, sympathy, and self-respect.

Prognosis and Diagnosis of Anti-Social Conduct.

It should become possible to diagnose delinquent or abnormal tendencies in children by the use of questionnaires standardised experimentally from results obtained from " normal," " abnormal," and delinquent children. Questionnaires eliciting such information would serve a double function. They would serve firstly as a prognostic social instrument designed for the protection of society by the early detection of social misconduct and anti-social tendencies, and secondly they would hold out hope for more effective remedial treatment in respect of anti-social conduct, since the chances of remedying anti-social behaviour are normally greater the earlier in life it is detected.

Psychological Insight gained from exhaustive Study of Children's Answers.

Including both papers in both Questionnaires, each pupil was asked a total of seventy-two questions of one kind or another. With a total of 256 pupils, this gave a grand total of 9,216 answers (excluding those of the experimental period). These answers contained extensive examples of the child's rationalisations, " conventionalisations," hypocrisies, moral realism, wishes, desires and interests. They afforded psychological insight into, and induced a sympathetic understanding of, the child's point of view. This is important inasmuch as we must know what the child's predominant interests, desires, hopes, fears and prejudices are, if we wish to appeal to him effectively.

BIBLIOGRAPHY

ABEL, THEODORA M.: "Moral judgment among subnormals" (*J. Abnor. and Social Psychology*, 1941, 36, 378-391).
ALBIG, WILLIAM: *Public Opinion*, 1939.
ALLPORT, GORDON W.: " Attitudes " (*Handbook of Social Psychology*, Ed. Carl Murchison, Ch. XVII, 1935).
BARKLEY, KEY L.: " Development of the Moral Judgment of College Students " (*Character and Personality*, Vol. X, No. 3, Mar. 1942, 199-212).
BRESLAW, BERNARD J.: " The Development of a Socio-Economic Attitude " (*Archives of Psychology*, No. 226, 1938).
BROWN, WILLIAM: " The Psychological Basis of Ethics (*Character and Personality*, Vol. VII, No. 1, Sept., 1938, 1-13).
BURT, CYRIL: *The Young Delinquent*, 1948.
CARRITT, EDGAR, F.: *The Theory of Morals*, 1935.
CHARTERS, WERRETT W.: *The Teaching of Ideals*, 1935.
GLUECK, SHELDON AND ELEANOR: *Criminal Careers in Retrospect*, 1943.
HADFIELD, J.: " Psychology and Morals—An Analysis of Character " (*International J. of Ethics*, Vol. XXXIV, 1923-24).
HARTSHORNE, HUGH and MAY, MARK, A.: *Studies in Deceit*, 1928.
HOPKINS, KENNETH D.: " Punishment in Schools " (*Br. J. of Educ. Psych.*, Vol. IX, 1939, 8-28).
MACKENZIE, JOHN S.: *A Manual of Ethics*, 1935.
McKOWN, H. C.: *Character Education*.
MALINOWSKI, BRONISLAW: *Crime and Custom in Savage Society*.
MORETON, D. E.: " Attitudes to Religion among Adolescents and Adults " (*Br. J. of Educ. Psych.*, Vol. XIV, June 1944, 69-79).
MULLINS, CLAUD: *Crime and Psychology*, 1943.
OLSON, WILLARD C.: " Character Education " (*Encyc. of Educ. Research*, 1941, 119-125).
PIAGET, JEAN: *The Moral Judgment of the Child*, 1932.
RAUBENHEIMER, ALBERT S.: " An Experimental Study of Some Behaviour Traits of the Potentially Delinquent Boy " (*Psych. Monographs*, Vol. XXXIV, No. 159, 1925, 106).
ROBERTSON, A.: *Morals in World History*.
STAGNER, ROSS: " Attitudes " (*Encyc. of Educ. Res.*, 1941, 69-75).
THURSTONE, LOUIS L.: *The Nature of Intelligence*.
URBAN, WILBUR M.: *Fundamentals of Ethics*, 1931.

An Investigation into the Effects of Punishment and Reward on Spelling.[1]

BY ANDREW GOLD.

(*Summary of thesis submitted in part fulfilment of the requirements for the Degree of Bachelor of Education at Glasgow University*, 1949.)

The main purpose of this work was to study the effects of punishment in contrast to reward on the process of learning in children. The experimental situation chosen for investigation was the learning of a spelling list.

The subjects were chosen from three primary schools and divided into three experimental groups : P, a punishment group ; R, a reward group ; and C, a control group ; each containing 40 children. The homogeneity of groups was established not on the basis of mental or scholastic age, but on spelling age. Inglis' Spelling test was administered to 300 subjects and the three parallel groups were selected with spelling ages ranging from 9 to 12.

A list of spelling words was tested on a class of subjects with a high spelling age. The most difficult words were removed, and thus a test of 50 words was made up.

The Experiment.—The list was written on the blackboard, and each group allowed 40 minutes to learn the words. The words were read aloud by the tester and repeated by the subjects, who then spelled them aloud. Then each individual was left to learn in his own way. The learning was completed before the morning interval. After the interval the children were asked to write the words on a sheet especially prepared. Without previous notice, the test was repeated on the following day, in order to isolate the effects of recency on retention. Two months later the test was given under ordinary conditions of a normal day's work.

Creation of atmosphere of punishment and reward.—Even before the list of words was given to the punishment and reward groups, a certain atmosphere was created. The punishment group received a severe reproof regarding the previous poor quality of their spelling, and were threatened detention on the following Friday and asked to do a punishment exercise. There were also threats of the loss of drill and games period and even of physical punishment. On the other hand, the subjects of the reward group were praised for their previous spelling abilities, and given incentives such as more free time on the following Friday, an extra drill or games period and a sweet. In the control group, the learning of the words was normal in the normal manner, and no such threats or promises were made.

Observations during learning.—The experimenter made some very useful observations during the learning period. In the punishment group it was noticed that a sort of concern existed and diminished until five minutes before the end, when very few expressions of fear could be observed. In the reward group the subjects started off in an atmosphere of excitement and enthusiasm, which lasted not more than 15 minutes, then the normal conditions prevailed. In the control group the normal conditions prevailed all the time.

RESULTS

	Test I			Test II			Test III		
	P	R	C	P	R	C	P	R	C
Mean	46·08	44·15	42·68	46·13	44·20	42·80	43·10	41·50	39·28
S.D.	3·32	4·40	5·84	3·38	4·33	6·01	3·22	4·45	7·35
S.E.	·53	·70	·92	·53	·69	·95	·51	·70	1·16

---

[1] Summarised by Z. Barbu.

As shown in the table, the mean of group P. is greater than those of groups R and C. The lowest S.D. and S.E. appear in group P., while the highest are in group C.

More illustrative is the comparison of these groups. The Critical Ratio between groups P and C is over 3, therefore, significant, while between groups R and C, is very far from being significant. The conclusion is that punishment proved a satisfactory stimulant for learning. Speaking from an educational viewpoint, the pupils required to be " pushed " from time to time.

The fact that reward did not prove to be an incentive to learning is explained by some considerations on the " modern child " to whom an extra drill and games period, or a sweet, do not appeal enough. The Critical Ratio between the P and R groups is 2·99, and therefore significant on the 0·01 level. There is no significant difference between the sexes.

Effects of Reward and Punishment on Learning as opposed to Performance.—A second experiment was set up, to establish the effects of punishment and reward on performance, as distinct from learning. This research aimed at an experimental re-examination of the thesis held by many psychologists—in the field of animal psychology mainly—that motives and incentives affect only the performance of what is already acquired, not the actual learning.

A test made up of 30 arithmetical problems was given to 106 fourth and sixth-grade girls. Four groups were formed : (1) A group publicly reproved for its mistakes ; (2) a group publicly praised for its performance ; (3) a group which, while hearing the praise and reproof administered to other groups, was ignored ; and (4) a control group.

Results.—On a first trial—before the administration of praise and reproof—the results of all groups were the same : 12 problems in 15 minutes. On the second day, both the praised and reproved groups gave an average performance of 16 problems, while the other groups showed no improvement. From this point, the performance of the first two groups started to decrease. A comparison shows that the Critical Ratio between the praise and control groups was 2·89 ; between reproof and control 2·05 ; and between reproof and praise 0·83. Thus the praise group showed an improvement over the control group significant on the 0·01 level, and the reproof group on the 0·05 level. The experimenter felt unsatisfied, however, and a repeat of the experiment showed no significant differences. From this the conclusion was drawn that the incentives of punishment and reward might be of little value where testing of performance alone was concerned, and could be effective where actual learning by the experimental subjects was involved. From a theoretical viewpoint it is assumed that learning and performance are, in fact, inter-connected and that sharp distinction between them seems to be artificial.

Historical : Experiments on Reward and Punishment.—Thorndike's experiment with chicks demonstrated that "reward strengthened, but punishment did not weaken the tendencies to which it was attached." From this the role of punishment in learning was more clearly understood. Nuttin's experiments marked another step in the psychology of learning. According to his research, the task-tension is a more important factor than success or reward and punishment in the process of learning. The process of learning increases if the learning situation (Gestalt) is so created as to leave the subjects in a state of tension of an uncompleted task. From a comparison between Nuttin's experiments and the learning of the spelling list it is pointed out that punishment is a more powerful factor in learning than task-tension.

The Psychology of Punishment and Reward.—Pickford has stressed the close association between punishment and reward and the law of success, in the sense that punishment is the reproof of unsuccessful reactions, and reward the praising of successful reactions. Regarding the problem from the viewpoint of depth psychology, Berg has pointed out that punishment can be the expression of the exploitation by the super-ego of aggressive components of the *id.*

From the educational point of view punishment and reward must be founded in the law of effect. In this sense, one can use pleasure or reward as a means of stamping in bonds of learning, and unpleasure or punishment as a method of destroying bonds. In the present research punishment appeared to have produced better effects than reward, but the explanation of this might involve reasons not easy to analyse here.

BIBILOGRAPHY

BERG, C. : " The Psychology of Punishment " (*Brit. J. Med. Psych.*, XX, 295-313, 1945).
HOPKINS, K. D. : " Punishment in Schools " (*Brit. J. Educ. Psych.*, IX, 8-28, 1939).
MCGEOCH, J.A. : *The Psychology of Human Learning* (1942).
MUNN, N. L. : *The Fundamentals of Human Adjustment* (1946).
NUTTIN, J. : " Respective Effectiveness of Success and Task-Tension in Learning " (*Brit. J. Psych.*, XXXVIII, 49-56, 1947).
PICKFORD, R. W. : "Psychological Aspects of Punishment" (*Int. J. Ethics*, LVIII, 1-17, 1947).
SCHONELL, F. J. : *Essentials in Teaching and Testing Spelling* (1942).
THORNDIKE, E. L. : *Psychology of Wants, Interests and Attitudes* (1934).
THORNDIKE, E. L. : *Man and His Works* (1943).
THOULESS, R. H. : *General and Social Psychology* (1937).

Schoolgirls' Attitudes to Films, Youth Clubs, Homework, Discipline and Sport.

BY MARGARET McGHEE.

(*Summary of thesis submitted in part fulfilment of the requirements of the Bachelor of Education Degree at Glasgow University*, 1948.)

This investigation had two aims :

(1) To *compile* an attitude test dealing with five aspects of adolescent schoolgirl life, viz. : Films, Youth Clubs, Homework, Discipline and Sport.

(2) To *administer* this test to secondary schoolgirls (ages varying from 12 to 17) and to apply statistical analysis to the results.

With the aim of producing a reasonably objective measuring instrument, a modification of Thurstone's psycho-physical method of equal appearing intervals was used to construct a suitable attitude scale. The initial statements were drawn up from the anonymous comments of children, of varying ages, on their attitude to the five topics. From these a list of statements—as concise and unambiguous as possible and varying from very favourable to very unfavourable attitudes—was compiled for each topic. With regard to Homework and Discipline, the children were not necessarily expected to *like* either, but provision was made for any who might, in addition to those who might not like but could tolerate both as necessary evils. In the scale for Homework, attitudes varied in intensity from (*a*) Homework should be abolished completely, to (*b*) Homework is necessary for success in life, or again to (*c*) a great satisfaction and pleasure comes from doing Homework.

The five groups of statements were presented to 55 university students of psychology, who were asked to give each statement an independent numerical rating, ranging from +5 to −5 ; 0 being assigned to neutral statements. From this was obtained an average positive or negative value for each statement. In addition, the eleven items in each group with the lowest S.D. were chosen and arranged in descending order. The intervals between the values of the five resulting scales varied considerably but the scale was more objective than an *a priori* scale.

Application of the Scales.—The statements, rearranged within each group, in random order, were presented to each of 185 senior secondary pupils (chosen to give representative variation in age and scholastic ability). The test was done anonymously. Date of birth and class required. The child was asked to put a tick (√) in each section, beside the item with which she agreed most of all and to put a cross(×) beside the item with which she disagreed most of all. The values from the students' ratings were inserted for each statement with a tick or a cross. Using these data a numerical measure or score (either positive or negative) was obtained for each Subject's Attitude to each of the five topics.

From statistical analyses made resulted the following conclusions:

Conclusions.

A. (1) When inter-correlated the attitudes to the five topics seem to be quite unrelated.

(2) The attitude scores to the five topics seem to be grouped into very high, neutral or very low, the intermediate scores rarely being achieved. (A possible explanation may be a tendency in young people to choose extremes and to be incapable of finer and subtler distinctions.)

(3) The general mean attitude to all five topics was positive and favourable—the attitude to Clubs being the most favourable of all.

B. (4) The correlation between age and attitude was very small and not significant. There was a general tendency towards a group which approved, a group which was neutral, and a group which disapproved, at each age level.

(5) The older girls have a more favourable attitude to the films, whereas Clubs have a greater appeal to the younger children. However, the personal and social elements in Clubs seem to make them much more popular than the Films at all ages and at all levels of I.Q.

(6) The group deviating most from the general mean was the 14-year-old group, which was rather negative to Films and very negative to Discipline.

(7) The mean of the 17-year-old group coincided four times out of five with the general mean, indicating a more balanced and moderate attitude.

C. (8) The correlation between Intelligence (as judged from raw scores of an adapted form of Raven's test) and attitudes yielded two significant correlations:

(*a*) Between Films and Intelligence, $-\cdot 71$, and
(*b*) Between Homework and Intelligence, $-\cdot 30$.

There was no significant correlation between any others.

(9) The most negative score towards Homework was achieved by those with the highest intelligence.

(10) The most negative score to *Discipline* was achieved by those of lowest intelligence.

(11) The most positive attitude to Films was achieved by those of lowest intelligence.

(12) Attitudes to Clubs and Sport seemed to be very favourable at all ages and at all levels of intelligence.

BOOK REVIEWS

The Nature of the General, Verbal and Spatial Bi-factors: F. SWINEFORD. (Chicago, Ill.: University of Chicago Press, Supplementary Educational Monographs, No. 67, 1948, pp. xi+71, price $2.00.)

This monograph gives a clear account of the application of Holzinger's technique of group-factor analysis to test scores of 952 children in Grades V-IX of Chicago schools. The battery was designed to yield a general factor and verbal and spatial group-factors, and therefore naturally did so. The factor pattern was highly consistent in the two sexes and in different grades, showing no tendency for greater differentiation of group-factors at later ages. It is interesting that spatial ability (when *g* is held constant) showed no appreciable growth with age. Perhaps the most original finding is that, though the *v*-factor tended to correlate with school marks in verbal subjects, and the space factor with Art and Mechanical Drawing, the correlations were so tiny. But to the reviewer it would appear that the reason why school marks depended so predominantly on *g* in this research is because the *g* tests themselves were largely verbal and arithmetical. It would be quite possible to re-analyse the correlations so that these tests, as well as the purposely verbal ones, would show considerable *v*-saturations, and such a *v*-factor might well be more educationally significant. Apparently the decision as to the best factor solution of a set of correlations is still largely subjective ; neither centroid, or bi-factor, or any other technique overcomes this difficulty. P.E.V.

Modern Trends in Psychological Medicine: Edited by NOEL G. HARRIS. (Butterworth and Co., Ltd., pp. xii-450, 50s.)

This is a comprehensive and valuable volume with articles by nineteen specialists. Of special interest to readers of this JOURNAL are the following : " The Psychology of Emotions " (S. Wright) ; " The Importance of Constitutional Factors " (N. G. Harris) ; " Marriage and Family Life " (E. F. Griffith) ; " The Future of Child Guidance " (W. Moodie) ; " Principles of Mental Hygiene " (H. V. Dicks); " Character Formation in Relation to Education " (H. R. Hamley).

Dr. Moodie's article on " Child Guidance " is very general in form : well adapted for the psychiatrist or educational reader who is largely ignorant of the Child Guidance Clinic and of findings as to its value, but otherwise lacking in precise evidence, or reasoned argument. Professor Hamley's article on " Character Formation " includes brief paragraphs on Conflict, Reconciliation, Faith, Democracy and Totalitarianism. But it is strange to see him quote a modern child psychiatrist to support the view that there is innate in human nature an urge to make sacrifices for loved people—a view which McDougall and others emphasised long ago. But perhaps Professor Hamley thought that a reference to psychiatrists would appeal more to the medical readers of this book rather than reference to McDougall, of whom many will have heard criticisms, though so few seem to have read any of his books.

Child Psychiatry: LEO KANNER, M.D. (Blackwell, pp. xxiv-752, price not stated.)

This second edition of the book published in 1935 has been entirely re-written and enlarged by over one-third of the original length. It is now probably the most comprehensive book (by a single author) on the subject. The treatment is practical and factual, rather than theoretical or dogmatic. Dr. Kanner is cautious and critical even when he finds definite value in some method of investigation or of treatment. For example, though he gives a section to the sympathetic discussion of play investigation and play therapy, he is very critical of Melanie Klein's fanciful speculations and interpretations.

As the book covers such a vast field of psycholog [illegible] topics (including intelligence, attitudes, projection methods, [illegible] every variety of psycho-somatic problem), the discussions of the individual topics—more especially the psychological—are often very brief ; but the volume would be a most valuable reference book for child psychologists working in a clinic, as well as to psychiatrists.

The Oxford publishers are too dignified to stamp the price on the book, but, as the first edition was 27s., this is probably at least 35s.

Hypnotism To-day : L. M. LECRON, B.A., and J. BORDEAUX, B.A., M.A. (with a foreword by M. H. Erickson, M.D.). (Heinemann, ix-278, price 25s.)

This is an entirely admirable book. It is by two non-medical psychologists, but it has the warm commendation of a leading research psychiatrist, so that is all right ! Indeed, two paragraphs from Dr. Erickson's introduction almost provide a good review of the book. " The authors," he writes, " are two psychologists, earnestly and honestly interested in furthering a scientific knowledge of hypnosis. Their purpose is to present to the lay and professional public simply and adequately the available information on hypnosis and to portray their own use of hypnosis as a psychotherapeutic and investigative procedure as demonstrated by their own experience under medical supervision. They have selected their material well, their exposition of it is excellent, and their presentation of their own studies is unassuming and straightforward . . . those who have worked extensively in hypnosis will feel grateful to the authors for offering a book to the public that presents hypnosis as a worthy, serious field of scientific endeavour."

The reviewer endorses that entirely. Two things must be added. The theoretical aspects of hypnotism and its relation to suggestion are thoroughly discussed. Also the manner of, and occasions for, using hypnotism in psychotherapy are equally well treated. One important point emphasised is that the usual criticism of hypnosis as a means of cure, namely that the results are only temporary, often applies also to other forms of psychological treatment, unless they are frequently repeated.

It is notable that these two consulting psychologists have actually read a good deal, besides ' practising.'

Statistics in School : W. L. SUMNER. (Blackwell, Oxford, pp. vi+183, 9s. 6d. net.)

This is a very handy book for more advanced studer [illegible] education, and for research workers ; and one or two [illegible] " Marking ") could profitably be read by students in training. The exposition is clear, and there are ample tables and graphs.

Fundamental Statistics in Pyschology and Education : J. P. GUILFORD. (McGraw-Hill Book Co., 2nd edit., 1950, pp. xiii+633.)

The second edition of this comprehensive book has been considerably extended and largely rewritten. More attention has been paid, in particular, to sampling statistics, analysis of variance, prediction, reliability and validity of tests, and scaling procedures. A very useful feature is the series of tables in the appendix, including tables on squares and square roots of numbers 1 to 1,000, significant coefficients of correlation and t ratios, chi square, F ratio, and Fisher's Z for different values of r. The book would be a useful addition to the library of any psychological laboratory. But as so often in American books, references to the work of some important English statistical psychologists are, unfortunately, lacking.

Problems of Social Policy: R. M. TITMUS. (His Majesty's Stationery Office and Longmans, Green and Co., pp. 596, 25s. net.)

This volume in the official series of United Kingdom Civil Histories, is chiefly concerned with the social consequencies of air attacks, the work of the hospital services, and the evacuation of mothers and children. Readers of this *Journal* are likely to find, especially in the treatment of the last of these topics, a great deal to interest them. Perhaps the discussion is at times verbose, but certainly the volume will be an invaluable reference book for facts.

OTHER PUBLICATIONS RECEIVED

The mention of a book in this list does not preclude a subsequent review.

Psychology : J. A. C. BROWN (Paul Elek, pp. 141, 8s. 6d.).

Correspondence Education in Australia and New Zealand : Compiled by S. A. RAYNER (Melbourne Univ. Press, pp. 119, 8s.).

Law and the Modern Mind : JEROME FRANK (Stevens, pp. xxxi-368, 25s.).

The Desert People : JOSEPH, SPICER and CHESKY (Cambridge Univ. Press, pp. xvii-288, 48s.).

The Inner Experience of a Psycho-analyst: THEODOR REIK (Allen and Unwin, pp. xiv-514, 21s.).

Culture and Personality : Edited by S. STANSFELD SARGENT (Viking Fund, pp. vi-219, 9s.).

A Girl Grows Up : RUTH FEDDER (McGraw-Hill, pp. 271, 16s.).

Situational Factors in Leadership : JOHN K. HEMPHILL (Ohio State Univ. Press, pp. 136, 15s.).

Educational Acceleration: Appraisals and Basic Problems (Ohio State Univ. Press, xii-153, 15s.).

Understanding Ourselves : MARY MACAULAY (Routledge and Kegan Paul, pp. xi-148, 7s. 6d.).

I Planted a Seed : Inspired by WILFRID BELL (Forbes, Robertson, Ltd., pp. 53, 4s. 6d.).

Mysterious Marriage : E. GRAHAM HOWE (Faber and Faber, Ltd., pp. 320, 15s.).

Fundamentals of Psycho-analysis : FRANZ ALEXANDER (Allen and Unwin, xi-312, 16s.).

Intelligence Testing and National Reconstruction : C. M. BHATIA (Hind Kitabs, Ltd., Bombay, pp. 83, Rs. 2).

The Educational Development of the Child: W. GLASSY and E. J. WEEKS (Univ. of London Press, pp. xiii-248, 8s. 6d.).

Everyday Problems of Childhood : E. MILDRED NEVILL (William Earl and Co., Ltd., pp. 156, 6s.).

The Explanation of Life : STEPHEN TH. BORNEMISZA (Rascher Edition, Zurich, pp. 272, no price given).

The Group Approach to Leadership-Testing : HENRY HARRIS, M.D. (Routledge and Kegan Paul, pp. x-288, 21s.).

The Foundations of Common Sense : NATHAN ISAACS (Routledge and Kegan Paul, pp. xii-208, 15s.).

Personality Maladjustments and Mental Hygiene : J. E. WALLACE WALLIN (McGraw-Hill, xiv-581, 30s.).

Problems of Social Policy : RICHARD M. TITMUSS (H.M. Stationery Office and Longmans, Green, pp. xi-496, 25s.).

Immediate and Retention Effects of Interpolated Rest Periods on Learning Performance : BERTRAM EPSTEIN (Teachers' College Columbia University Contributions to Education, No. 949, xiii-77, 12s. 6d.).

The Teacher in a Planned Society : R. W. RICH (Univ. of London Press, pp. 120, 5s.).

Nai Talim Sevagram : Hindustani Talimi Sangh (pp. not numbered, Rs. 3).

Rural Education : A. B. ALLEN (Allman and Son, Ltd., pp. 255, 7s. 6d.).

Leçons de Pedagogie : ROGER COUSINET (Presses Universitaires de France; pp. xi-261, 400 francs).

Le Personnalisme : EMMANUEL MOUNIER (Pr. Univ. de Fr., pp. 136, no price given).

Propos sur l'Esthétique : ALAIN (Pr. Univ. de Fr., pp. 118, 180 francs).

La Névrose d'Abandon : GERMAINE GUEX (Pr. Univ. de Fr., pp. 140, 300 francs).

La Divinité, Étude Psychanalytique : PAUL DIEL (Pr. Univ. de Fr., pp. 218, 460 francs).

Introduction a l'Épistémologie Génétique : (Tome I—La Pensée Mathematique) : JEAN PIAGET (Pr. Univ. de Fr., pp. 361, 700 francs).

Introduction a l'Épistémologie Génétique: (Tome II—La Pensée Physique) : JEAN PIAGET (Pr. Univ. de Fr., pp. 355, 700 francs).

VOLUME XX. NOVEMBER 1950 PART 3.

POST-GRADUATE TRAINING OF TEACHERS IN PSYCHOLOGY.

BY P. E. VERNON
(Professor of Educational Psychology Institute of Education, University of London).

I.—Part-time courses in psychology for teachers. II.—Reasons for success of teachers with higher degrees in education : (1) *selection* ; (2) *maturity* ; (3) *bias of courses.* III.—*Possible improvements.*

A DISCUSSION on Post-graduate Training in Psychology was held at the Reading meeting of the British Psychological Society, on April 1st, 1950. Professor Rex Knight introduced the topic, and was followed by Dr. N. A. B. Wilson. As the third speaker, I dealt with the kind of psychological training in which I have most experience, and the following represents my personal views—not the views of the British Psychological Society, nor, of course, of the Universities of Glasgow and London, at which I have taught.

I.—PART-TIME COURSES IN PSYCHOLOGY FOR TEACHERS.

Most teachers graduate at a university in some subject or subjects other than psychology, and during their teacher training they receive only an elementary psychological course. Either directly after they have qualified, or else after some years of teaching, they may return to the university for post-graduate training leading, usually, to the M.A. in Education, or to the Ed.B. or B.Ed. degrees in Scotland. Such training is mostly part-time ; the students carry on their teaching work simultaneously, and study psychology only in the evenings, at week-ends and during holidays. Edinburgh, however, insists on a full-time course, and some other universities have other schemes, about which I must plead ignorance.

Compared with the training of students who take honours degrees in psychology, these part-time courses are very incomplete. And my only reason for mentioning these students in this symposium is that they actually constitute one of the main sources of professional psychologists in Britain. Roughly one hundred men and women acquire higher degrees in education every year, apart from those who specialize in the history or theory of education or in comparative education rather than in psychology. Of these hundred, some stay in teaching ; some become training college or W.E.A. lecturers, though few that I can think of have become university lecturers. A considerable proportion go into administration. But it is they who provide most of the psychologists at Child Guidance Clinics and the educational psychologists employed by education authorities, often without scarcely any further training except on the job. Just about half of the members of the Committee of Professional Psychologists (Mental Health) have M.A. or Bachelor degrees in education. Among Scottish members the proportion is far higher. If we add to the English members those with other part-time degrees or diplomas in psychology, they greatly out-number the professional psychologists who possess full-time degrees. Moreover, these students appear to be capable of adapting to a variety of other professional jobs, in the Fighting Services for example. In 1942 the psychology or personnel selection departments were set up in the Services by fully qualified psychologists, mostly from the National Institute of Industrial Psychology. But, thereafter, the technical staff was expanded chiefly by recruitment from this source. By 1945 approximately ten of the Service psychologists who had posts of major responsibility were ex-teachers, whose basic psychological training had been extremely limited. I am all the more impressed by the professional usefulness of the Glasgow Ed.B., because our students who took M.A. Honours in psychology so rarely became professional psychologists. I would suggest that, at least in the past, there has been a similar wastage at other British universities. A considerable proportion of psychology students has not proceeded to any post-graduate work and has been lost to the profession.

II.—REASONS FOR SUCCESS OF TEACHERS WITH HIGHER DEGREES IN EDUCATION.

Why then do students with higher degrees in education turn out well? Three main reasons may be suggested.

(1) *Selection.*—First, the students are highly selected. Most of them have honours degrees to begin with. Actually, London is introducing a method of entry open to ordinary degree and even to non-graduate teachers, though I think with adequate safeguards, In addition, they are self-selected in that they usually realise, or soon come to realise, the amount of work involved, particularly in carrying out a thesis investigation. Thus they enter only if they are quite strongly motivated. It may be said that the motive often consists less of a genuine interest in psychology than of a desire to escape from the rut of teaching, or to better themselves. Many of the women who wish to get into child guidance belong to what might be called the tender-minded type. At Glasgow, their reactions to any apparatus in the laboratory involving electrical circuits were not unlike their reactions to a mouse running across the floor. And, in spite of all my efforts to teach them statistics, they often absorbed one principle only, namely to go and ask the laboratory steward to show them. But I would not criticise such shortcomings, for I am convinced that these students are doing immense good to innumerable maladjusted and backward children. The more tough-minded (mostly men), even though their desire to enter administration may have been less praiseworthy, nevertheless, tackled statistics, laboratory work, psychoanalytic theories and Gestalt, with uniform efficiency. The important point, in my view, is that all these students had a sense of vocation which is often lacking in the university student who takes psychology in order to get a degree, and/or because he is interested in the subject.

(2) *Maturity.*—The next factor is the maturity of the students. Not only have they had considerable experience of handling people and of responsibility, but also this experience continues while they are undergoing their psychological training. For this reason I prefer the part-time method to Edinburgh's scheme of full-time training, often following immediately after the first degree, though I admit that Edinburgh achieves higher academic standards among its B.Ed.s than we at Glasgow did with the Ed.B.s. A particularly outstanding bunch of students were those who entered from the Services in 1946. They possessed *par excellence* this quality of mature experience.

One is led to wonder whether the psychology honours student who proceeds straight away to research for the Ph.D. (as I myself did) is not somewhat lacking in this respect. Too many professional psychologists in this country as well as in America, become too exclusively rat experts, or statisticians, or gadgeteers. In America, Thurstone discovered a zero correlation between interest in psychology and interest in people. It seems to me then that the Education degree method of combining post-graduate training with training in living may be one of the best ways of overcoming this autistic tendency among psychologists.

(3) *Bias of Courses.*—In spite of the inadequacies of the Ed.B. and M.A. courses, they have got a strongly applied bias. Perhaps they do succeed to some extent in helping students to approach professional psychological work along sound lines. Though they have not the time to acquire much fundamental knowledge, they get to know their way about the field of psychology, and where to go for more information on problems that they meet with. I have no doubt that the thesis investigation is the most valuable part of this. Not only does the prospect of tackling a genuine, even if tiny, piece of research, frighten off the less suitable half of students who would otherwise clutter up the course, but also the quality of the thesis generally correlates better with the students' final first or second class or fail result than do any of the theory papers; and it probably gives a better prediction of ultimate professional competence. It is, of course, difficult to mark theses because different students receive such varying amounts of help from their supervisors. But the competent student generally produces his own topic of investigation, and even if one does spend many hours in showing him what to do, one has all the more chance of gauging his qualities of mind. I am well aware of the dangers of the doctrine of formal discipline, but, if it were possible to get evidence, I should expect this research on some educational or psychometric problem to show a very high degree of transfer to any subsequent professional work. Better than any other method it helps students to put their academic knowledge and their reading into practice, and to show them that psychology is not merely a matter of books and lectures.

What does the formal psychological training amount to outside the thesis work? Students usually take a course in Education at the same time as educational psychology, so that their attendance at lectures and seminars is limited to two hours at the end of their working day for about two days a week (or else Saturday mornings), for—usually—two years. Even including the psychological part of their Teachers' Certificate, the total time for the average student extends only to 200 hours at Glasgow and 150 at the London University Institute of Education. In terms of 25 hour (i.e., 1 hour per week for 1 year) units, the distribution is roughly as follows:

	Glasgow Units.	London Units.
General, Social and Abnormal Psychology	2½	1½
Educational and Child Psychology, Tests and Statistics	2½	4
Laboratory Work	2½	0
Testing Demonstration and Practice	½	½
	8	6

I must apologise if I am maligning other universities which may have much more thorough courses; but Glasgow and London between them probably produce more than half the yearly total for the country. Not only is this class work about one-third as extensive as that of most psychology honours students, but also the reading, or the writing of essays if any, is often extremely meagre.

Personally, I am not greatly worried about the paucity of fundamental training in general psychology, though I do not want any less than 2 units. Provided the students can find their way around in general, social and abnormal psychological literature, I consider that they are little if at all handicapped in their professional careers because they do not know the first thing about figural after-effects, the goal-gradient hypothesis, the biosphere, or object-cathexes. am quite unrepentant also in not regretting greatly the entire absence of a laboratory course at London. My experience at Glasgow was that practice in the psychophysical methods, in measuring reaction times, and the like, had very little transfer value. Try as I would, I could not get the students to grasp the applicability of methods outside the particular context. The Method of Production was, for example, indissolubly linked with the Muller-Lyer illusion, the Method of Limits with phenomenal regression, and so on. I valued the laboratory work more because, in the absence of any time for seminars or for good discussions at lectures, they did provide some opportunity for getting to know the students as individuals. It is possible, however, to devise more realistic experiments on learning, mental tests, marking of essays, and even in social psychology. I would suggest that the traditional course should be scrapped and rebuilt along these lines.

III.—POSSIBLE IMPROVEMENTS.

Any time that can be given to the practical testing of children, to interviewing and getting case-histories, or to field work outside the laboratory seems to me of the utmost value. Obviously, one cannot hope to produce competent clinical, vocational, or other types of professional psychologist in ten to twenty hours ; but one can stop some of the leaks and point out some of the grosser errors of technique. Thus the main inadequacies appear to me to lie in the vocational part of the lecture course. I simply cannot cover in the available time the minimum essentials of :

> Child development, delinquency and maladjustment, backwardness, applications of psychology to school subjects, educational and vocational selection and guidance, description and uses of tests, test construction and item analysis, intelligence and heredity, school marking and examining, design of experiments, statistical methods, reliability and sampling, understanding the uses of factor analysis and analysis of variance, assessment of personality, interests and attitudes, etc., etc.

It seems hopeless also to try to introduce the project system or other more enlightened methods of teaching than the orthodox lecture. Nor can one begin to deal with the specific background needed for each type of professional work—in clinics, for education authorities, in the Services, and so on. It helps to give out duplicated notes which outline the main points in the above topics, but I am coming round to the opinion that I may need to adopt to some extent the American system of required reading. The difficulty with this is, of course, the lack of good and cheap books in these fields.

To sum up : There is no doubt that the present system of post-graduate training of teachers in psychology yields a rich source of professional psychologists. Yet at the same time, I am far from satisfied with the standards reached, at least at the universities with which I have been, or am, associated. The main steps that occur to me for improving this situation are :

(1) Still more stringent selection should be applied. Some kind of qualifying examination might be given to intending entrants to prove that they have a certain amount of background knowledge, that they are genuinely interested in psychology and are not merely careerists.

(2) The system of required reading and quizzes, however objectionable to educationists, is worth considering. But we need more books by British authors to cover adequately such fields as child development, test construction, intelligence and heredity, applications of psychology to classroom teaching (in actual, not ideal, schools), vocational psychology, etc.

(3) The initiative of Heads of Departments should be applied to the development both of more extensive field work (testing, controlled observation, interviewing, etc.) and of laboratory work which has a closer bearing on practical educational, vocational and social problems.

THE EMOTIONAL RESPONSES OF ADOLESCENT GROUPS TO CERTAIN FILMS

PART I.

BY W. D. WALL AND W. A. SIMSON.†
(Department of Research, University of Birmingham Institute of Education.)

I.—*Introduction—the films seen: Age and sex distribution of the samples.* II.—*General results—principal effects of the films as a whole—identifications—sadness, tears and fright—erotic stimulation—attitudes to others—differences between comedies and dramas.* III.—*Sex differences—sadness and tears—fright—identifications—effects of specific films.* IV.—*Conclusions and summary.*

I.—INTRODUCTORY.

IN previous papers,[1] attempts have been made to assess broadly the impact of the commercial cinema on the social and psychological development of boys and girls in the early teens. The present paper is concerned with a closer analysis of the effect of certain specific films upon samples of the boys and girls who saw them.

In one of our major enquiries, 2,069 adolescents from all types of school in the Midlands and between the ages of 13·0 and 16·11 were asked to name a film which they had seen in the previous fortnight and to answer a series of specific questions upon it (See Appendix).

By this means it was hoped that representative and sufficiently large samples of adolescents would answer questions on a limited number of films which they had seen under the ordinary circumstances of their habitual attendance. In fact, the spread of interest over the films available was wider than had been anticipated, which meant that the number of samples sufficiently large to be reasonably reliable was small.[2] Twelve films only produced replies from groups numbering thirty-five or more. These films, however, are fairly representative of the whole body of films screened in the given fortnight, with the exception that there was no horror film among them.

The method of gathering the information is open to the disadvantage that there was no means of ensuring comparability of sampling from film to film. But in this it reflects a real situation. Not merely were the subjects of the enquiry ignorant, at the time of seeing the film, of their forthcoming participation in this research and, therefore, somewhat less liable to sophistication, but it is also to be supposed that any exercise of preference in the choice of film would be unhampered. It was felt that the situation in which a party of adolescents is taken to see a film or films in the knowledge that they will be questioned about it afterwards, would lose by artificiality what it gained in representativeness of sampling.

---

† The material on which this paper is based has been prepared for publication by the writers. The gathering of the material, its statistical analysis and the discussion of the results has been the work of a small group of teachers of which the authors are Chairman and Secretary, respectively.

[1] This *Journal*, Vol. XIX, Pt. 1, Pt. II; *Educational Review*, Vol. 1, Pt. 1, Pt. 2; and *Education*, No. 1948.

[2] 293 films were mentioned. For the distribution of attendances see "The Adolescent and the Cinema," *Educational Review*, Vol. 1, Pts. 1 and 2, and "The Film Choices of Adolescents" this *Journal*, Vol. XIX, Pt. 2.

The films included four comedies (*London Town, Time of their Lives, Monsieur Beaucaire, Kid from Brooklyn*), four which might be classified as dramas (*Odd Man Out, Root of All Evil, The Green Years, Razor's Edge*), one musical biography (*The Magic Bow*), one thriller (*Send for Paul Temple*) and two rather different westerns (*The Outlaw* and *Smoky*).

In all, 476 boys and 379 girls answered questions on one or other of these films. Table 1 shows the age and sex distribution of the total sample, film by film.

TABLE I

Age and Sex Distribution of the Sample, Film by Film
(476 Boys and 379 Girls.)

Age Group.	Boys.				Girls.				Total.	
	13+	14+	15+	16+	13+	14+	15+	16+	B.	G.
Film.										
London Town	24	25	15	3	29	30	6	6	67	71
Time of their Lives	31	30	18	17	13	5	5	0	96	23
Monsieur Beaucaire	2	8	15	12	6	3	2	6	37	17
Kid from Brooklyn	5	3	5	7	4	9	5	0	20	18
Odd Man Out	4	15	13	10	7	16	13	9	42	45
Root of all Evil	1	3	11	6	1	7	11	4	21	23
Green Years	5	1	5	7	19	14	10	2	18	45
Razor's Edge	1	3	12	8	10	11	9	2	24	32
Magic Bow	8	6	13	10	7	21	11	16	37	55
Paul Temple	12	7	6	4	9	0	0	0	29	9
Outlaw	0	7	12	19	3	4	2	0	38	9
Smoky	21	9	14	3	2	17	10	3	47	32
Totals	114	117	139	106	110	137	84	48	476	379

It will be noticed that, with a few exceptions, mainly in the smaller samples, the age distributions for boys and girls are reasonably comparable. Where this is clearly not the case, as for example in the groups who saw *Monsieur Beaucaire,* it is possible that the film held a special attraction for adolescents of a particular age and sex. The overall numbers in the different age groups suggest that, on the whole, the sample of boys is fairly evenly spread over the age range with rather more 15-year-olds than might have been expected. Among the girls there are notably few 16-year-olds and a relatively high proportion of 14-year-olds. Just over half the girls (55 per cent.) and considerably less than half the boys (39 per cent.) were attending or had attended grammar schools. These anomalies of sampling should be borne in mind in considering the figures and conclusions which follow.

II.—The general results.

The questions were framed to elicit responses which might throw light on the effects of the chosen film upon different areas of the emotional life. In Table II, the analysis is set out in five main categories according to whether the question was aimed at discovering (i) immediate emotional effects—laughter, tears, fright and so on ; (ii) relatively enduring emotional effects of sadness, happiness, fright and nocturnal and day dreams ; (iii) effects upon attitudes to others, especially on heterosexual relationships ; (iv) possible identifications with characters in the film ; (v) the effects upon some aspects of the viewer's scheme of values.

TABLE II—PROPORTIONS OF THE VIEWING GROUPS WHO MADE POSITIVE RESPONSES TO VARIOUS QUESTIONS ON SPECIFIC FILMS (476 Boys ; 379 Girls ; 12 Films).

	COMEDIES								DRAMAS								MUS. BIO		THRIL'R		WESTERNS					
	London Town		Time of theirlives		Mons. Beaucaire		Kid from Brooklyn		Odd Man Out		Root of All Evil		Green Years		Razor's Edge		Magic Bow		Paul Temple		Outlaw		Smoky		Mean Per cent.	
N	B. 67 %	G. 71 %	B. 96 %	G. 23 %	B. 37 %	G. 17 %	B. 20 %	G. 18 %	B. 42 %	G. 45 %	B. 21 %	G. 23 %	B. 18 %	G. 45 %	B. 24 %	G. 32 %	B. 37 %	G. 55 %	B. 29 %	G. 9 %	B. 38 %	G. 9 %	B. 47 %	G. 32 %	B. 476 %	G. 379 %
I.—IMMEDIATE EMOTIONAL EFFECTS :																										
Q.1 Laughed	98	99	100	100	100	100	100	100	52	60	85	78	100	82	61	67	83	96	59	56	87	100	66	65	83	83
Q.2 Felt like crying	17	28	**1**	**13**	0	0	0	19	**45**	**74**	**19**	**52**	**65**	**98**	**26**	**59**	**21**	**83**	7	11	8	44	**49**	**78**	21	47
Q.3 Cried	3	9	**0**	**9**	0	0	0	0	**10**	**49**	**0**	**22**	**17**	**31**	4	6	**0**	**53**	0	0	0	33	**9**	**35**	4	21
Q.4 Frightened	2	4	**6**	**26**	0	0	0	0	14	24	0	13	6	11	9	16	0	2	48	67	16	33	13	0	**9**	**16**
Q.5 Bored	22	17	8	9	8	0	10	6	14	9	33	9	6	7	39	19	24	13	11	0	32	22	11	17	18	11
Mean per cent.	28	31	23	31	22	20	22	25	27	42	28	35	39	46	29	33	26	49	25	27	29	47	29	39	26	35
II.—LASTING EMOTIONAL EFFECTS :																										
Q. 6 Sad afterwards	12	8	6	9	0	0	5	0	53	74	14	17	53	60	26	41	**3**	**26**	0	0	3	44	40	52	18	28
Q. 7 Happy afterwards	60	87	88	96	78	94	75	86	17	18	67	65	41	67	26	28	72	92	34	56	39	44	55	65	54	67
Q. 8 Frightened	9	3	3	9	0	0	0	0	5	7	0	0	0	11	0	3	0	0	**31**	**44**	3	0	4	4	4	7
Q.18 Dreamed	18	17	12	26	11	18	10	6	2	4	5	26	35	24	4	6	**3**	**43**	17	11	5	33	**16**	**22**	15	20
Q.19 Remembered	62	58	64	52	49	76	55	69	62	74	57	52	71	49	·60	63	**69**	**90**	31	44	71	56	**40**	**74**	68	53
Q.21 Enjoyed	80	72	81	87	73	82	80	94	60	65	71	61	65	82	57	66	80	96	66	56	63	56	62	52	70	72
Mean per cent.	40	57	42	46	35	45	38	43	33	40	36	37	44	49	29	34	38	58	30	35	37	39	36	45	36	44
III.—EFFECTS ON ATTITUDE TO OTHERS:																										
Q.12 More loving to Boys	12	20	6	**0**	11	12	0	13	5	**4**	**0**	17	18	36	13	13	7	47	35	11	5	22	7	13	10	17
Q.14 More loving to Girls	21	21	**26**	9	43	0	30	6	**31**	2	29	**30**	53	18	13	3	42	2	17	22	63	11	24	4	33	12
Q.10 More loving to Parents	22	33	8	4	0	6	0	6	18	4	14	17	**18**	**56**	4	3	10	19	3	0	8	22	20	4	11	15
Q.11 Help others	24 -	36	**18**	**39**	14	0	0	13	41	53	29	35	29	51	52	**31**	17	36	21	11	11	22	36	35	24	30
Mean per cent.	20	27	14	13	17	4	8	9	24	17	18	25	30 .	40	20	13	19	31	19	11	22	19	22	14	19	19
IV.—IDENTIFICATION :																										
Q.22 Character	25	31	26	35	33	29	10	25	22	20	38	44	47	47	39	22	**35**	**60**	**55**	**11**	63	44	27	39	35	34
V.—VALUES AND ATTITUDES :																										
Q. 9 Long for money	**0**	**17**	5	0	16	18	10	0	7	0	19	9	0	7	9	3	0	9	4	11	5	0	0	13	6	7
Q.15 Feel life dull	38	35	37	30	33	18	30	25	**36**	**9**	29	30	35	33	30	34	21	17	41	11	55	78	38	30	35	29
Q.17 Long to travel	23	30	9	13	38	29	20	25	17	16	19	13	6	16	70	75	59	74	11	0	29	78	49	48	29	35
Q.13 Bad successful	**10**	**30**	18	9	16	12	10	6	19	24	29	30	24	16	13	19	10	11	14	33	45	33	**20**	**0**	19	19
Q.16 Good successful	60	68	76	74	57	41	50	56	49	31	81	66	**94**	**60**	35	66	65	79	69	89	58	67	**84**	**57**	65	63
Q.20 Shocked	**8**	**21**	12	22	5	12	10	6	19	20	19	22	24	11	39	25	10	17	21	0	34	33	25	9	19	17
Q.23 True to life	38	42	11	4	11	0	5	22	70	65	71	61	77	91	52	81	69	66	41	56	37	44	64	67	45	50
Mean per cent.	25	35	24	22	25	19	19	20	31	24	38	33	37	33	35	43	33	39	29	29	38	48	40	32	31	31
Mean Overall per cent.	29	34	27	29	26	24	22	25	29	31	32	33	38	42	30	33	30	46	28	26	34	40	33	34	30	33

All percentages are rounded to whole numbers. The means given marginally and horizontally have been calculated from the unrounded figures. Figures in bold type indicate significant (·05 or better) sex differences based on values of Chi-square derived from the raw numbers and using Yates' correction where necessary.

Many of the questions allowed for gradations of response but since, in most cases, the numbers are too small to warrant fine analysis, all positive responses have been grouped together: negative, indifferent and unanswered were also scored together. In spite of the smallness of the numbers, the proportions have been shewn as percentages to allow of rapid comparison. The sex differences indicated in bold type in the table, however, are based on a calculation of the value of chi-square (d.f.1) from the raw numbers.

It is recognised that to some unknown extent there may be a tendency towards a positive response, in spite of the random arrangement of the alternative answers. Hence marginally are shewn the average percentages of positive replies to each question and horizontally the averages for each film and for each of the five groups of questions. The impact of the various films on these sample audiences may perhaps best be gauged by comparing the percentages of positive responses either with the mean for the question or with the overall means of 30 per cent. positive responses for boys and 33 per cent. for girls.

One further caution is necessary before certain general trends are discussed. It cannot be assumed that because, for example, a high proportion of boys or girls state that they felt like crying at a particular film, the emotion felt was, therefore, intense. The figures of Table II merely indicate the spread within the group of a particular emotional response. They indicate not that girls, for example, felt more or less deeply about this or that, but that a smaller or greater proportion acknowledge feeling a particular emotion.

With this in mind we may examine Table II for the light it throws upon statements which are often made concerning the effects of films on young people. If we may judge by the present sample, it is clear that laughter and a relatively enduring happiness or satisfaction are the principal effects. It seems obvious also that the general trend of the moral teaching of the films, for what it is worth, is seen by adolescent audiences to be that the good win in the end. The stimulus which the screen gives to the fantasy life, too, may be inferred from the proportions who remembered, specially enjoyed or dreamed about something in the film under discussion. About one-third of all the adolescents in the sample also claimed to identify with one or other of the characters portrayed. Experiences of sadness, both immediate and enduring, though significantly more common among the girls as would have been expected, are not absent among the boys. Small proportions of boys (mean 4 per cent.) and larger numbers of girls (mean 21 per cent.) confess to actual tears. On the other hand, frightening experiences seem to have been few and confined principally to the one thriller.

It is often alleged that the cinema makes its audience dissatisfied, that it makes life seem dull and that it stimulates a desire for money as a means of escape. With wide variations from film to film, it appears that of the present sample, about one-third thought that their film experience made life seem dull and about the same proportion experienced a longing for travel. Less than 10 per cent. on the average found themselves longing for more money.

It is also alleged that, with its emphasis on the relationships between the sexes, the cinema stimulates and organises the erotic impulses of young people. The evidence provided by the boys and girls who saw these films, gives some support to this contention, especially in the case of the boys of whom on the average only a little less than one third claim that the film which they saw made them feel more loving towards their girl friends. Considerably fewer girls made a similar admission. It will be noticed, however, that there is considerable variation, especially among girls, from film to film in this respect. Such films as the *Magic Bow* and the *Green Years*, where the love story is a central theme, affect substantial proportions of both sexes.[1] On the other hand a thriller like *Send for Paul Temple* or the story of Somerset Maugham's *Razor's Edge* affect relatively few.

The differences between the broad categories of response are of considerable interest. Fewest admit to any effect of their cinema experiences on their *attitudes to others* (mean 19 per cent. both for boys and girls). Of the films concerned, *Odd Man Out* had as a theme certain aspects of social injustice and the *Razor's Edge* dealt with semi-christian ideas of altruism and renunciation in the person of Larry. It is, therefore, significant that these two films produced the greatest proportions of positive replies to the question, " Did it make you feel that you would like to help other people ? " In the *Green Years, London Town* and *Smoky* there were more or less altruistic themes, expressed, however, in highly personal terms, and these also produced a proportion of positive responses to the same question. Conversely the emptiness of such films as *Monsieur Beaucaire* and *Kid from Brooklyn* is reflected in the replies. Only two of the films dealt with the theme of parent-child relationships—*London Town* and the *Green Years*—and in neither was it greatly or strikingly stressed. Some, especially of the girls, appear to have been affected by this.

Many of the questions which dealt with *lasting emotional effects* produced relatively high mean percentages. It should be remembered that the subjects answered the questionnaire on a film seen in the preceding fortnight. Hence the length of the " memory " involved might have been as little as two or three days and would not have been more than three weeks ; for most it would have been about seven to ten days. The principal lasting effects seem to have been enjoyment and happiness. Experiences of sadness and fright seem with many to have faded. In the figures for " dreaming " and " remembering " (Q. 18 and 19) there is some evidence of continuing fantasy on which more light will be thrown when we come to consider the qualitative responses to certain questions.

The remaining three categories, with considerable variations from question to question, are on the average closer to the general mean.

Among the films themselves there is considerable variety in the responses they provoke. *Kid from Brooklyn,* a rather commonplace musical comedy about a milkman turned boxer, contrasts markedly with *The Green Years,* Cronin's story of the orphan boy who wins through to become a doctor. *Monsieur Beaucaire*, a farcical and ruthless adaptation of Booth Tarkington's novel, about a barber masquerading as a duke, and *Send for Paul Temple,* a thriller in which the amateur detective assists Scotland Yard to unmask the leader of a band of diamond thieves, appear to have left the few girls who saw them relatively unmoved. *The Magic Bow,* on the other hand, with its romanticised biography of Paganini and its love theme, provoked a high proportion of positive responses from the girls who saw it.

The four comedies almost universally provoked laughter and left behind them a feeling of happiness and enjoyment. There was little in them to make their audiences feel sad or cry, and apart from the serio-comic ghosts in *Time of their Lives,* little to cause fright. With the exception of *London Town,* with its setting on the London variety stage, they are not considered by any large proportion of the sample audiences to be true to life. *The Magic Bow* provoked a pattern of responses very similar to that of the comedies which, in fact, it closely resembles. There are, however, four striking differences. Very much higher proportions of those who saw it, probably because of its pseudo-biographical content, thought it true to life. The pathos of the film too seems to have affected considerable proportions of the girls. It is also noteworthy that a higher percentage of the girls who saw it, wished to be someone in the film (usually Phyllis Calvert) than for any other. The comedies generally apparently had little effect upon the girls' feelings towards boys: *The Magic Bow,* however, affected nearly half the girls who saw it.

---

[1] All the films were seen by at least one adult observer who reported on a specially prepared schedule. This has been supplemented where necessary by additional information from the summaries and criticisms contained in the relevant *British Film Institute Bulletin.*

The groups of film classified as 'drama' dealt more or less realistically with human behaviour and, in particular, with various aspects of human loyalty and helpfulness. The greater realism is reflected in the high proportions of their audiences who thought them true to life. They were enjoyed by the majority, though by markedly smaller proportions than the comedies. They were more provocative of emotional response in general, especially of feelings of sadness, both immediate and lasting. This is especially true of *Odd Man Out* and *The Green Years.* The happy ending of the latter, however, appears to have done its work; less than 20 per cent. of those who saw *Odd Man Out,* which ends in tragedy, as compared with more than half those who saw *The Green Years,* stated that they were happy afterwards. In contrast to the comedies generally and to most of the other films, these four dramas had a widespread influence on attitudes to others and in particular they appear to have aroused the desire to help other people.

The thriller, *Send for Paul Temple,* is chiefly remarkable for the relatively high proportions of those who admit to fright both at the time of seeing the film and afterwards. In no other film do more than an insignificant proportion of the audience admit to enduring fright and no more than a quarter to immediate fright. Of those who saw *Send for Paul Temple,* more than half admit to having been frightened during the film and rather less than half consider that the fright lasted for some time afterwards.

Of the two films classified as 'Westerns,' neither is quite typical of the *genre.* In *The Outlaw,* two 'bad men' carry out a vendetta over a beautiful half-caste, Rio (Jane Russell). In *Smoky* the principal theme is the adventures of a beautiful horse and his devoted master. In neither, apparently, was there anything that the boys found at all frightening. The girls particularly, and nearly half the boys found *Smoky* provocative of sadness which for some lasted afterwards. Possibly, as we shall see, because of some rather unfortunate publicity, rather more than seventy per cent. of the boys who saw *The Outlaw* remembered various impressions afterwards, and among all the twelve films it was the one which led most to think that the bad are successful. It also appeared to offer to many of the boys who saw it a fantasy identification with the successful bad man. Not unconnected with this is the fact that nearly two thirds of the boys who saw it—the highest proportion for any of the films studied—stated that it made them more loving towards their girl friends.

III.—Sex differences.

Of the sex differences shewn in Table II, the most consistent is that shewn in response to Question 2, in respect to seven out of the twelve films and in response to Question 3 in respect to six of the films. Where there is anything to provoke it, more girls than boys feel like crying and actually cry at what they see on the screen. A similar, though less striking trend (significant statistically in one case only) is that shewn in response to Question 6 more girls than boys experience enduring sadness.

This tendency for girls to experience sadness is, of course, in accordance with what is known of emotional differences between the sexes and confirms a similar finding from different data.[1] Its extensiveness among the girls and especially in response to *The Green Years, Magic Bow, Smoky* and *Odd Man Out,* is remarkable—the expression, " there was not a dry eye in the house " seems less of an exaggeration than might have been anticipated !

Most of the other sex differences in the table are to be explained in terms of the content of the particular films. In *London Town* for example, one of the main themes is the way in which the devotion of a daughter enabled her struggling father to achieve success on the variety stage. There is the somewhat conventional picture of poverty and of an undeserving character who is at first successful until the daughter plays a trick upon him to give her father his chance. The part played by the daughter offered identifications for the girls in the audience and probably accounts for the higher proportions of girls who felt that the film made them long for more money and that the bad were successful. Those who were shocked by something seen in the film referred almost invariably to the lowness and shortness of some of the dresses worn by the chorus girls in the picture.

About a quarter of the girls who saw *Time of their Lives* reacted with fright, a feeling which it appears did not endure. From the replies to Question 20 it seems that startlement at some of the ghostly antics shewn in the film rather than real fear was what was meant, though one or two mention the incident in which dead bodies are thrown down a well. It is of interest to note that the execution scene in the film, which some adult observers thought not altogether suitable for children or adolescents, was not, in fact, mentioned by anyone.

There is little stress on sex themes in this film though there are a male and a female ghost. The differences in reply to Questions 12 and 14 are probably a reflection of the boys' general tendency to reply positively to Question 14. That a relatively high proportion of girls felt the film affected their desire to help others may be a reflection of the treatment (admittedly trivial) of one of the themes of the film—the attempt of the ghosts to clear their memory of the imputation of treachery.

Odd Man Out was a very different film portraying the last days of a leader of an illegal political organisation and his devoted sweetheart. In many ways it appears to have profoundly moved its adolescent audience and to have been markedly different in its effects from other films. More boys than girls found the exciting incidents of the film made them think their own life rather dull—though not more than many of the other films. Similarly the loyalty of Kathleen, the heroine, to Johnny McQueen may account for the differences shewn in response to Question 14.

Apart from the effects of sadness already commented on, there are two quite striking sex differences shewn in the responses of boys and girls to *The Green Years.* The film appears to have offered identifications with the hero, Roddy, and the heroine, Kitty. Roddy is an orphan cared for by relatives and, in particular, by an old reprobate of a grandfather. His is a story of success through hard work and his reward is, among other things, the love of Kitty. The girl has parents who are devoted to her, kindly and sympathetic. It is apparent that for boys and girls there is a difference in the stress of the film—for the boys on the success which rewards effort and less on parent-child relationships, for the girls more on the parent-child situation and on the whole less on the success though Kitty has ambitions to become a great singer. The differences between boys and girls in the proportions answering positively to Questions 10 and 16 seem to reflect these aspects of the film's handling of its theme.

---

[1] *Educational Review, loc. cit.,* Part 2, p. 122.

The *Magic Bow* had two sympathetic female parts ; it exploited period costume, music and romance. Its appeal for the girls seems to have been very great and largely accounts for the sex differences observed in Table II. It seems to have provoked more girls to dream than any other film and to have been specially remembered by an overwhelming majority. The heroine and the second female lead between them seem to have offered identifications to a higher than average proportion of the girls who saw it.

In the latter respect, *Send for Paul Temple* provides a contrast. The central figure of the play is the detective. The main female lead, Steve, is little more than a lay figure. The fact, therefore, that half the boys wished to be the detective, whereas only one girl thought she would like to be " Steve, as she is a girl detective," is not unexpected.

Smoky appealed more to the girls who saw it than to the boys if we may judge from the replies to Question 19. Largely, it seems this is because of the portrayal of the horse which is one of the central themes of the film. On the other hand, boys in general appear to have put a somewhat different construction upon the moral values of the film. Right does triumph in the end in the person of Clint and the bad brother Frank is defeated. More boys than girls thought that the film made them believe that things turn out all right for good people in the end, but also—probably because Frank seems to get his brother into difficulties and is not harshly dealt with—a substantial proportion also thought that bad people triumph. It is perhaps legitimate to infer from this and from the qualitative evidence derived from Questions 18, 19, 20 and 21 that whereas the boys saw the film mainly in terms of a story about cowboys, and the Wild West, the girls were more inclined to respond to it as a tale in which the horse played the leading part.

IV.—CONCLUSIONS.

From the evidence so far discussed, certain broad general conclusions emerge. The great variation from film to film in the proportions responding positively to any particular question—a range of from 0 to 98 per cent. in the case of Question 2 (*felt like crying*) and almost as extreme in Questions 4 (*fright*), 6 (*lasting sadness*), 7 (*happiness*), 15 (*feel life dull*), 17 (*desire to travel*) and 23 (*true to life*)—points to the fact that any particular film is not accepted merely as a few hours' passive entertainment or idle pastime. There is evidence of considerable difference in the responses provoked by different films, and generalizations about the effects of the cinema must give way to an analysis of the effects of particular films or specific themes.

There is evidence too that even the light and superficial treatment of a particular theme, for example that of parent-child relationships, when it allows a fairly close emotional participation, does influence attitudes. Similarly the treatment of good and evil characters has an effect that is quite striking. When success clearly rewards virtue, especially in direct romantic terms as it does in *The Green Years*, the result is unequivocal—that the good win in the end seems a general opinion. Conversely where, with a greater approach to the subtlety of real life, it is left in some doubt as in *Odd Man Out* and *The Razor's Edge* doubts seem to arise in many minds, and where there is marked confusion of values, as in *The Outlaw*, this too appears to be reflected in the results.

The answers to Question 23—" Would you say that the film was true to life ? "—suggest that adolescents are able, certainly wherever they have experience or knowledge to guide them, to distinguish realism from reality. This is not to say that many are not frequently misled nor that realism lavished upon a theme which is false to true æsthetic or human values would not mislead many more. It emphasises the hopefulness of the suggestion made elsewhere[1] that growing boys and girls should be taught through informed discussion to appraise the worthwhileness of what they see and that steps should be taken to provide alternative experience in real human situations in which they may learn some of the many things of which at present the screen is the only teacher.

Table II also provides evidence in a general way of how sex differences and probably idiosyncracies of individual experience and development operate to determine which among the many stimuli offered will be accepted by any child as relevant to his or her own emotional growth. On this, the most crucial of all the problems connected with the Cinema, the qualitative and detailed information provided by Questions 18 to 22 throw some further light. The proportions answering positively to Questions 19 (" Did anything in the film remain in your thoughts for a long time? "), 22 (" Would you like to be any characters in the film? ") and 21 (" Did you specially enjoy anything in the film? ") are high enough to suggest that, for very many, particular incidents, themes or characters may become the starting point for reverie or be dwelt on for days or even weeks as a pleasureable emotional stimulus.

From Table II there is little evidence that these films made their adolescent audience markedly dissatisfied with their lot. Few were made to long for lots of money; more, but rarely more than one third, felt their own lives rather dull—a common enough adolescent feeling even in the days before the Cinema—and a number, depending quite noticeably upon the setting of the film, wanted to travel.

It is possible that the questions asked did not get to the heart of the matter, that they were too vague and general or that there were serious omissions. To some extent at least the possibility is covered by the space allowed for the qualitative and freer responses an analysis of which is presented in Part II of this paper. One cannot, however, even at this stage, escape the conclusion that only a few of the many films shewn and seen make a lasting impression or are likely radically to modify the attitude of an adolescent audience.

On the other hand our method can give no measure of the cumulative effect of a biased or partial view of life presented in a sequence of undistinguished films. That attitudes are at least temporarily modified even by relatively slight and unreal films seems to be apparent. We have little means of knowing how far this represents a permanent modification.

SUMMARY.

1.—An analysis of the responses of 476 boys and 379 girls aged 13·0—16·11 to twelve films is presented. These responses were obtained after an interval of from two days to under three weeks.

2.—The principal responses seem to have been amusement and enduring satisfaction.

3.—Among the boys fairly constant proportions tend to state that the film seen made them more loving to their girl friends. Among the girls, however, the proportions who are made more loving to their boy friends are directly affected by the nature of the film.

4.—Experiences of sadness both immediate and relatively enduring, are reported by substantial numbers of girls and by many more girls than boys.

5.—There is some evidence that films provide material for continuing fantasy.

---

[1] This *Journal*, XIX, *loc. cit.*

6.—Films dealing more or less seriously with human themes have a generally more wide-spread effect upon the emotional life than do comedies and have more effect upon attitudes to others.

7.—Identifications with principal characters or stars in the films appear to be made by about one-third of adolescent audiences. Where, however, the film portrays a character, particularly sympathetic for one sex (e.g., for girls *Magic Bow*; for boys, *Paul Temple*) more than half the group tend to say that they would like to be a person in the film. Certain films (*Green Years* is an outstanding example) offer different identifications and different interpretations to the two sexes.

8.—There is evidence of discrimination between films which are true to life and those which are not.

9.—There is little evidence that many young people are made to feel their own lives dull by their cinema experiences. Fewer still are made to long for lots of money.

10.—Experiences of fright, particularly of enduring fright, appear to be relatively rare. In the main, in the present series of films, they occur only among those who saw the detective film. An incident which adult viewers thought might be frightening was not commented on by a single adolescent.

APPENDIX.

The following is the form of questionnaire given in the major survey from which the samples discussed above were drawn. Only the questions relevant to the present paper are given. The responses to other questions have been analysed elsewhere.

This investigation is being carried out by people who do not know you, but they are particularly anxious that you should give your own personal and frank opinion. Please be absolutely honest. If you feel you cannot, please hand this paper in blank.

Do not write your name anywhere on this paper.

Boy. Girl. School..

Put a ring round your age last birthday: 12 13 14 15 16 17

Choose a film you have seen at any time (*except Saturday morning or Sunday*) in the *last fortnight*, January 20th to February 1st.

Write its title here..

Where did you see it? ....

*** Please answer ALL these questions about the film by underlining your answer among those given at the side:

1.—Did this film make you laugh? ..	Not at all. A little. Very much. Quite a lot.
2.—Did it make you feel like crying?	Very much. A little. Quite a lot. Not at all.
3.—Did it make you cry?	Quite a lot. Very much. Not at all. A little.
4.—Did it frighten you?	Very much. A little. Quite a lot. Not at all.
5.—Did it bore you?	Not at all. A little. Very much. Quite a lot.
6.—Did the film make you feel sad afterwards?	For a little while. Not at all. A long time.

7.—Did it make you feel happy afterwards ?	For a long time. Not at all. A little while.
8.—Did it make you feel frightened afterwards ?	For a little while. A long time. Not at all.
9.—Did it make you long for lots of money ?	Yes. Didn't think of it that way. No.
10.—Did it make you feel more loving towards your parents ?	Didn't think of it that way. No. Yes.
11.—Did it make you feel you would like to help other people ?	No. Yes. Didn't think of it that way.
12.—Did it make you feel more loving towards your boy friend(s) ?	No. Didn't think of it that way. Yes.
13.—Did it make you feel that even bad people are successful in this world ?	Didn't think of it that way. No. Yes.
14.—Did it make you feel more loving towards your girl friend(s) ? .	Yes. No. Didn't think of it that way.
15.—Did it make you feel that your own life is rather dull ?	No. Didn't think of it that way. Yes.
16.—Did it make you feel that things come out all right for good people in the end ?	No. Yes. Didn't think of it that way.
17.—Did it make you want to travel ?	Didn't think of it that way. No. Yes.
18.—Did it make you dream about it afterwards ?	Yes. No.
If " Yes," what was it that made you dream ?	....
19.—Did anything in the film remain in your thoughts for a long time ?	No. Yes.
If " Yes," what ?	....
20.—Were you shocked by anything in the film ?	Yes. No.
If " Yes," what ?	....
21.—Did you specially enjoy anything in the film ?	No. Yes.
If " Yes," what ?	....
22.—Would you like to be any character in the film ?	Yes. No.
Can you say which one, and why ?	....
23.—Would you say that the film was true to life ?	No. Yes.

THE SOCIAL AND EDUCATIONAL BACKGROUND OF EMERGENCY TRAINED TEACHERS AND REASONS FOR THEIR CHOICE OF THE PROFESSION.

By PETER J. T. GRANT
(*Assistant Education Officer, Cheltenham*).

I.—*Introduction.* II.—*The questionnaire.* III.—*Number of students involved in the enquiry.* IV.—*Questions and answers.* V.—*Comparison with previous investigations.* VI.—*Summary and conclusions.*

I.—INTRODUCTION.

THE Government Emergency Scheme for the recruitment and training of teachers has presented the country with an entirely new type of recruit to the teaching profession, necessitating new methods in tackling the problem of teacher selection and training. The scheme was launched in 1944 to meet the post-war shortage of teachers in England and Wales. It provided intensive courses for candidates over the age of twenty-one, who had been engaged in some form of national service during the war. Those of us who have been concerned with the training of these students have been impressed by the interest shown in their personal qualities by all manner of people.

Teachers, for instance, want to know more about the people who are to be their future colleagues; administrators and other educationists wish to learn more about their social attributes, their educational attainments and their ability to fit into the teaching profession, and the public wish to know what these men and women have to offer as teachers.

To attempt to answer these questions in a word would be an impossibility. I, therefore, prepared a Questionnaire (while I was a lecturer at an emergency training college), to cover a comprehensive range of questions, mainly concerning students' personal history and the circumstances leading to their decision to join the teaching profession.

II.—THE QUESTIONNAIRE.

So that we might have some common starting point, only students who were training for work in junior schools, were invited to answer the questionnaire. Men and women from four colleges were asked to answer the questions, which, in most cases, required only a tick at the appropriate answer. It was felt that the answers might be more accurate if the Questionnaire remained anonymous. At the end of the Questionnaire is a personal profile in which students were invited to assess their own personal qualities on a five-point scale.

III.—NUMBERS OF STUDENTS AND COLLEGES INVOLVED IN THE ENQUIRY.

Oakley Training College (Men), Cheltenham. 160 men.
Oakley Training College (Women), Cheltenham. 99 women.
Redlands Training College (Mixed), Bristol. 40 men and 7 women.
Wynyard Hall Training College (Women), Durham. 94 women.

VI.—QUESTIONS AND ANSWERS.

To conserve space in this account, many of the minor details have been grouped together under one heading. Questions 1–5 deal with personal matters of age and family commitments.

QUESTION 1.

Are you training for Infant or Junior Schools ?

QUESTION 2.

Indicate your age .... years.

Ages.	Men.	Women.
20/24	20	74
25/29	87	72
30/39	79	46
40/51	14	8

Observations : There is a wide age range for both men and women, although there are more women in the lower age ranges. Generally speaking, the weight of numbers with men lies between 25 and 36 years, and with women between 22 and 26 years.

QUESTION 3.

Sex M........... F...............

As previously indicated, the numbers were Male 200, Female 200.

QUESTION 4.

Married................ Single...............

	Men.	Women.
Married	146	18
Single	54	182

Observations : This question reveals an interesting feature of the scheme. Here we see some 70 per cent. of men married and of mature years, turning to a new career. Fewer women are married, yet even here the proportion is not inconsiderable, a factor consequent upon the clause in the 1944 Act which states that henceforth marriage will be no disqualification for employment as a teacher.

QUESTION 5.

Children : Girls .... Boys ....

Men 103 have families ranging from 1 to 7 in number.
Women 12 have families ranging from 1 to 2 in number.

Observations : From the male point of view at least the majority have their full share of family responsibilities while embarking upon a new career. They also have the advantage of first-hand experience with children.

Questions 6 to 10 seek to determine the academic background of these candidates.

QUESTION 6.

What type of school did you attend as a child ?

	Men.	Women.
(*a*) Secondary Grammar	108	120
(*b*) Central Selective	17	11
(*c*) Central Non-Selective	6	2
(*d*) Elementary	29	31
(*e*) Technical	21	10
(*f*) (i) Private Day with normal Leaving Age 16 and over	6	9
(ii) Private Day with normal Leaving Age below 16	4	7
(*g*) Boarding	9	10

Observations : The majority of both men and women attended grammar schools. Quite a large percentage did not attend schools beyond the elementary school stage.

QUESTION 7.

Did you like school ?

	Men.	Women.		Men.	Women
(*a*) Very much ..	52	97	(*b*) Much	60	43
(*c*) Fairly well ..	68	42	(*d*) Not very much ..	19	16
(*e*) Not at all ..	1	2			

Observations.—The majority of candidates of both sexes enjoyed their days at school, and there is a considerable definiteness about this.

QUESTION 8.

Was your attitude to school affected in any marked way by any of the following considerations :

	Men.	Women.
(*a*) School Discipline ..	51	74
(*b*) School Curriculum ..	53	91
(*c*) Personality of Teachers ..	134	145
(*d*) Methods of Teaching ..	48	45

Observations : Students' attitude to school was very largely affected by the personality of the teachers—whether good or bad : possibly an important factor in selecting candidates for the teaching profession.

QUESTION 9.

Did you gain a school Certificate before leaving school ?

	Men.	Women.
Yes	86	81
No	114	119

Observations : We must bear in mind that these are junior school candidates.

QUESTION 10.

At what age did you leave school ?

	13	14	15	16	17	18
Men	1	34	48	76	27	14
Women	1	39	35	79	35	11

Observations : The mean age of leaving school would appear to be 16.

Questions 11 to 13 deal with the choice of teaching as a career.

QUESTION 11.

When you were about to leave school did you give definite consideration to becoming a teacher ?

(*a*) Yes : Men 56, Women 82. (*b*) No : Men 114, Women 118.

Observations : Numbers reveal quite an interesting number having given thought to the idea of becoming a teacher at an early age. Let us now see what happened to these hopes.

QUESTION 12.

If (*a*), did any of the undermentioned considerations influence your choice :

(*a*) Influence of some other person	(i) Teacher	Men 8, Women 19.
	(ii) Parent	Men 5, Women 18.
	(iii) Relative	Men 4, Women 13.
	(iv) Friend	Men 2, Women 10.
(*b*) Interest in children		Men 29, Women 72.
(*c*) Natural desire to teach		Men 35, Women 47.
(*d*) Good conditions of work		Men 12, Women 6.
(*e*) Satisfactory pay and pensions scheme ..		Men 2, Women 5.
(*f*) Reasonable security		Men 18, Women 11.

Observations : Considerations affecting choice on the part of Men were (1) Natural desire to teach, and (2) Interest in young children. Women were influenced by (1) Interest in young children and (2) Influence of some other person, who was usually a teacher or parent. The influence of the teacher is also noticeably strongest with the men; perhaps a pointer towards the value of the teacher in vocational guidance for boys and girls about to leave school.

QUESTION 13.

Why did you not become a teacher then?	Men.	Women.
(a) Other employment more attractive	11	24
(b) Lack of money to pursue a course of training ..	59	107
(c) Objections from parents	6	16
(d) Prevented by national service	26	16
(e) Insufficient academic qualifications	25	42
(f) Unable to gain admission to a Training College	7	5

Observations: Candidates who wished to become teachers on leaving school were prevented from doing so mainly from economic reasons. To a lesser extent insufficient academic attainments were responsible and in the case of a fair proportion of men, the demands of national service provided the stumbling block.

QUESTION 14.

Previous employment before entering College.
Only occupations where five or more candidates were involved are listed.[1]

Men	(i) Clerks.	(ii) Salesmen.	(iii) Service (Regular).
Women........	(i) Clerks.	(ii) Civil Service.	(iii) G.P.O.

Observations: Most candidates have been trained for some occupation. The others are people, who, on leaving school, went into the forces for the duration of the war, at the conclusion of which they found themselves untrained for any occupation. The large number of candidates of both sexes from the clerical and associated occupations of the office category, is interesting. These are avenues of employment which require a person of considerable ability and reasonable education, but offer little opportunity for advancement, and the work itself is sometimes tedious and uninspiring for the more adventurous minds.

Questions 15 to 16 deal with candidates' efforts to study after leaving school.

QUESTION 15.

Since leaving school did you attend courses in further education arranged by:

	Men.	Women.
(a) Technical Schools or Colleges	62	76
(b) Evening Institutes	91	100
(c) W.E.A.	30	35
(d) Your employers	31	17
(e) H.M. Forces	89	38

Observations: Further education has been well patronised by candidates of both sexes. There would appear to be a genuine desire to improve their educational background. Considerable importance is attached to courses run by H.M. Forces.

---

[1] Other occupations were:

Men: Accountant, Artist (Commercial), Banker, Boot and Shoe Worker, Builder, Bus Driver, Carpenter, Chemist, Civil Servant, Commercial Traveller, Draughtsman, Engineer, Farmer, Insurance Agent, Mercantile Marine, Metal Craft Worker, Miner, Newspaper Production Worker, Own Business, Plumber, Policeman, Post Office Worker, Salesman, Sheet Metal Worker, Surveyor, Technical Worker.

Women: Accountant, Aircraft Worker, Banker, Beauty Culturist, Book-keeper, Church Worker, Comptometer Operator, Dental Attendant, Dressmaker, Drawing, Employment Officer, Farmer, Florist, Governess, G.P.O., Housewife, Hairdresser, Industrial Worker, Laboratory Assistant, Librarian, Local Government Clerk, School Matron, Supplementary Teacher, Telegraphist, Medical Orderly, Nurse, Nursery Nurse, Photographer, Saleswoman, Shop Assistant, Telephone Engineer, Tobacco Worker, War Service, Welfare Officer.

QUESTION 16.

State main subjects studied in Question 15[1]

Among *Men*, in order of importance, are: (1) English; (2) Mathematics; (3) Geography; (4) History. *Women* were interested in (1) Shorthand and Typing; (2) English; (3) Handicrafts; (4) Teaching; (5) Mathematics. There was a very wide range of subjects chosen by both men and women.

Observations: The availability of the subjects in further education has an important bearing on the choice of subject, but there is obviously a great demand for after school study of English and Mathematics. The writer feels that most students made a very real attempt to pursue these studies with interest, and that the knowledge acquired will prove valuable to them in the teaching profession.

Questions 17 to 25 deal with the choice of teaching as a profession.

QUESTION 17.

Did you decide to teach before or after the Government announcement of the Emergency Scheme?

Before: Men 28; Women 22
After: Men 172; Women 178

QUESTION 18.

Which of the following considerations influenced your choice?

	Men.	Women.
(*a*) Influence of individuals:		
(1) Teacher	36	23
(2) Parent	12	20
(3) Relative	22	21
(4) Friend	57	52
Group Total	127	116
(*b*) Growing interest in young children, through:	Men.	Women.
(1) Having a family of your own	48	10
(2) Membership of a youth club	48	42
(3) Church activities	43	49
(4) Your employment	28	48
(5) Voluntary work in child welfare, social welfare, etc.	19	22
Group Total	186	171
(*c*) Dislike of your employment, e.g.:	Men.	Women.
(1) Not interested in job	37	30
(2) Unsatisfactory conditions	27	12
(3) Poor prospects	24	16
(4) Poor pay	11	5
Group Total	99	63

---

[1] Subjects studied were: Accountancy, Agriculture, Art and Crafts, Architecture, Aviation, Biology, Building, Botany, Banking, Business Method, Chemistry, Civics, Commercial Subjects, Child Study, Cooking, Domestic Science, Choral Singing, Drama, Dancing, Discussion, Dressmaking, Economics, Electricity, Elocution, Esperanto, Engineering, English, Films, First-Aid, French, Fashion Design, Folk Dancing, Foreign Affairs, Farming, Gardening, General Knowledge, Geography, German, Greek, Handicrafts, History, Home-Nursing, Insurance, Latin, Logic, Law, Librarianship, Leather Work, Mathematics, Mechanics, Music, Navigation, Needlework, Nursing, Philosophy, Psychology, P. T., Printing, Public Administration, Public Speaking, Politics, Religion, Religious Leading, Salesmanship, Sheet Metal, Shorthand and Typing, Spanish, Speech, Science, Surveying, Social History, Sunday School, Tailoring, Teaching, Telegraphy, Textiles, Wireless, Woodwork, Youth Work, Zoology.

(d) Through advice contained in:	Men.	Women.
(1) Advertisements	6	15
(2) Books	18	11
(3) Newspaper articles	13	18
(4) Ministry Pamphlets	68	62
(5) Lecture on wireless on the subject of recruitment to teaching profession	9	16
Group Total	114	122

	Men.	Women.
(e) Attractive prospects of the profession	52	31
(f) Favourable financial terms during training	27	33

Observations : Among the men the most important consideration affecting choice of this new profession was (*d*) 4, i.e., Advice contained in Ministry Pamphlets; second, Influence of a friend; and third, Attractive prospects of the profession. For the women first choice also goes to the Ministry Pamphlets; second, Influence of a friend; and third, Growing interest in young children, through various contacts with them. This last result is interesting in its close correlation with the replies to Question 12. The important conclusion to be drawn from these results is that the majority of students of both sexes were motivated partly by a genuine interest in young children, which had arisen out of personal contact with them in various ways.

QUESTION 19.

Why did you decide to teach in a Junior School?	Men.	Women.
(a) Interest in young children	121	146
(b) Able to meet academic requirements	54	38
(c) Advice of a member of the teaching profession	43	46
(d) Variety of work rather than specialisation required	92	70

QUESTION 20.

Since making the decision, have you had an opportunity to teach?

Yes : Men, 118 ; Women, 134.
No : Men, 82 ; Women, 66.

QUESTION 21.

In what type of school was this teaching done?
Most students practised in junior or infant schools.

QUESTION 22.

Age and sex of children.

Similar to Question 21. It is interesting to note that over 20 men worked in infant schools.

QUESTION 23.

Main subjects taught.

Observations : Mainly general subjects.

QUESTION 24.

Did this experience reassure you of your ability to teach?

Yes : Men, 117 ; Women, 140.
No : Men, 1 ; Women, 4.

	Men.	Women.
It was (a) Very successful	31	18
(b) Successful	54	69
(c) Moderately successful	3	5
(d) Rather unsuccessful	0	1
(e) Complete failure.	0	0

Observations: The preliminary teaching experience for both men and women appears to have been invaluable both from the point of view of experience, and in confirming the decision to proceed with training for the profession.

QUESTION 25.

How much did this experience influence your decision to become a teacher in a junior school?

(*a*) Very much: Men 63, Women 87. (*b*) Much: Men 25, Women 29.
(*c*) Somewhat: Men 17, Women 14. (*d*) A little: Men 3, Women 4.
(*e*) Not at all: Men 10, Women 6.

Observations: This question is complementary to the previous one, and the replies reinforce those of that question.

Questions 26 and 27 are sociological, and deal with candidates' non-academic interests.

QUESTION 26.

State any special hobbies or interests not directly connected with teaching, e.g., philately, rock-climbing, chess, etc.

There were 99 hobbies and interests listed by men and women. Men chose, in order of preference, games, hiking and camping, gardening, and chess; while women preferred games, needlework, hiking and camping, and reading.[1]

QUESTION 27.

Have you been associated with children's organisations other than school?

Cubs	Men 24;	Women 14.	Scouts ..	Men 41;	Women 3.
Guides ..	Men —;	Women 53.	Brownies ..	Men —;	Women 32.
Boys' Brigade	Men 11;	Women 4.	Sunday School	Men 45;	Women 109.
A.T.C. ..	Men 25;	Women —.	Kindergarten	Men —;	Women 9.
G.T.C. ..	Men —;	Women 9.	Army Cadets	Men 5;	Women —.
Youth Club ..	Men 69;	Women 58.	Sea Cadets	Men 2;	Women —.

Observations: Both men and women appear to be most frequently connected with Youth Clubs and Sunday School.[2]

The last question asked "If, after some years in the teaching profession, you lose interest or "get tired" of it, would you resign and return to your former employment or some other employment?"

Men ..	Yes 121;	No 34;	Don't know 45.
Women ..	Yes 134;	No 40;	Don't know 26.

---

[1] Hobbies mentioned were: Archery, Archæology, Architecture, Art, Astronomy, Athletics, Aviation, Badminton, Ballet, Bee-keeping, Bible Study, Birds, Boating, Books, Botany, Boxing, Bridge, Campanology, Carpentry, Carving, Chess, Choirs, Church, Cinema, Club Work, Collecting Books, Collecting Records, Committee Work, Concerts, Correspondence, Country Dancing, Countryside, Crafts, Crocheting, Crosswords, Cycling, Dancing, Domestice Science, Dogs, Discussion, Drama, Draughts, Dressmaking, Esoteric Truth, Esperanto, Farming, Fencing, Film Projecting, First-Aid, Fishing, Flying, Food and Drink, Folk Dancing, French, Furniture, Gardening, Gliding, Hiking and Camping, Historical Ass., Housewifery, International Club, Knitting, Leatherwork, Local Government, Men, Model-making, Motoring, Music, Modern Languages, Needlework, Numismatics, Nursing, Opera, Painting, Philately, Philosophy, Photography, Piano, Picnics, Poetry, Politics, Poultry, Psychology, P.T., Radio, Reading, Riding, Rock-climbing, Rowing, Sailing, Scouting, Singing, Sight-seeing, Skating, Ski-ing, Speech, Social Work, Swimming, Table Tennis, Theatre, Toy-making, Travel, Violin, Watch and Clock Repairing, Wild Flowers, Woodwork, Writing, Youth Clubs, Youth Hostels.

[2] Others mentioned were: Remand Home, Church Dramatic Club, Sports Organisations, Boys' Choirs, Campaigners, P.T. Clubs, Evening Institutes, Youth Fellowship, Boys' Orphanage, Girls' Friendly Society, Ambulance, Sea Rangers, Girls' Guildry, Young Farmers' Club, Girls' Life Brigade, Red Cross Cadets.

THE PERSONAL PROFILE.

Candidates were invited to rate certain of their personal qualities on a five-point scale from A to E. " A " indicating very great possession of the quality, and " E " very little. The results were as follow :

		E	D	C	B	A
General Health Rating	Men	0	1	22	133	44
	Women	0	1	32	116	51
General Intelligence	Men	0	1	102	92	5
	Women	0	1	125	74	0
Verbal Ability	Men	0	4	100	80	16
	Women	0	12	120	58	10
Sense of Humour	Men	0	0	51	117	32
	Women	0	1	50	107	42
Speed of Work	Men	1	20	119	48	12
	Women	4	14	134	42	6
Practical Ability	Men	2	14	100	68	16
	Women	2	12	106	78	2
Artistic Ability	Men	18	42	82	46	12
	Women	14	44	93	43	6
Industry	Men	0	0	86	100	14
	Women	2	10	106	78	4
Reliability	Men	0	4	42	120	34
	Women	0	2	64	122	12
Sociability	Men	0	4	54	116	22
	Women	0	8	60	110	—

Observations : It is a matter of conjecture whether any great degree of reliance can be placed upon self-assessment of this type. Personal qualities which appear to require a high rating in the teaching profession are General Health, Intelligence, Sense of Humour, Industry, Reliability, Sociability. It is probably true that persons selected for admission to a course of training were selected on their possessing these qualities to a degree. Some qualities are, of course, more objectively assessed in oneself and these it will be observed are more normally distributed. These include artistic ability, verbal ability and practical ability. There are no noticeable sex differences in these ratings. The evidence suggests that the assessment of personal traits ought to play some part in the selection of candidates for the teaching profession.

V.—COMPARISON WITH PREVIOUS INVESTIGATIONS.

Professor C. W. Valentine[1] found that for university candidates the chief motives for becoming teachers were :

Men (1) To be in a profession with one's favourite studies.
(2) Economic desirability which also included the Board of Education grant.
Women (1) Liking for teaching.
(2) Fondness for children.

He also found that there was no truth in the view that the advice of secondary school heads was the chief influence in deciding pupils to choose the teaching profession. W. B. Tudhope[2] discovered in the case of normal training college students that the chief motives for men were (1) attraction of a secure job ; (2) fondness for children. Women were of the same opinion. He goes on to say that motive (3) was interest in a special subject ; (4) desire to continue one's own education ; (5) fondness for teaching ; (6) the possibility of doing good, and the good salary. Some weight is carried amongst both men and women by a parent's wish. The present investigation dealing with emergency trained candidates finds the following chief motives for wishing to enter the teaching profession :

---

[1] " Reasons for choice of the Teaching Profession by University Students "—C. W. VALENTINE, this *Journal*, Vol. IV, Pt. 3., November, 1934.

[2] " Motives for the choice of the Teaching Profession by Training College Students "—B. TUDHOPE, this *Journal*, Vol. XIV, Pt. 3., Nov., 1944.

(*a*) Candidates who decided that they would like to take up teaching when they were about to leave school did so because of :

Men (1) Natural desire to teach.
(2) Interest in children.
Women (1) Interest in children.
(2) Desire to teach.

At this age both men and women admit that they were influenced most by (*a*) Teacher ; (*b*) Parent.

(*b*) Candidates who decided definitely to enter the teaching profession at a later date did so because of the following reasons :

Men Growing interest in young children and young people through
- (*a*) Having a family of their own.
- (*b*) Membership of a youth club.
- (*c*) Church activities.

Women Growing interest in young children through
- (*a*) Church activities.
- (*b*) Membership of youth clubs or employment.

Both men and women were agreed that the most important source of persuasion was the Ministry of Education pamphlets on the subject of recruitment for the teaching profession. Friends, it appears, were the most important personal advisers at this stage.

Negative factors like " unsatisfactory conditions of work " were the least important. " Attractive prospects " and " favourable financial terms during training," so vital to the university and normal training college candidate, were, in this investigation, considered of little import.

VI.—SUMMARY AND CONCLUSIONS.

Persons subject to enquiry.—Men and women in their late twenties or early thirties, who after the war, took the opportunity of training for the teaching profession. The majority of men and some women were married and had children.

Educational background.—Candidates have attended all types of schools, mainly grammar. About 15 per cent. left school at the age of fourteen, 40 per cent. are in possession of a school certificate. In the main they enjoyed their time at school ; the personality of the teacher, they say, was the main influence for good or ill.

Desire to teach.—A few thought of becoming teachers at the time of leaving school, influenced by a " natural desire to teach " or because of their " interest in children." They were prevented from achieving their objective at that time, mainly by lack of money.

Former employment.—Candidates are recruited from a wide range of employment. Clerical work predominates in the case of both men and women.

Further education.—There has been a keen demand for further education by most candidates. The choice of subjects is very extensive.

Decision to teach.—Most candidates decided on teaching after the announcement of the Emergency Scheme. Advice on this step came chiefly from Ministry of Education pamphlets and the personal influence of friends, while the urge to teach came from a "a growing interest in children through contact with them," in various ways.

Pre-College training.—Pre-College teaching experience is considered to have been well worth while.

Hobbies and interests.—Interest in hobbies and young people's organisations is an important part of the social and intellectual background of these candidates and an important pointer to their suitability for the teaching profession.

The final question indicates that most candidates would be prepared to leave the profession if, after a few years, they lost interest or otherwise found themselves unfitted for it.

Personal profile.—The personal profile is an attempt at self-assessment of personal qualities. Its success or otherwise can only be judged by comparing individual assessments with the recorded judgment of tutors and others intimately concerned with the candidates.

This investigation reveals to us a type of entrant to the teaching profession, mature in years, with much experience of the art of living. There are many occasions when the academic background is weak, but everywhere there is evidence of the pursuit of knowledge and in nearly every case of the acquisition of non-academic experience and skill which will prove of great value to the practising teacher. The type of person who has been selected by the entrance boards is generally a person of good character and more than average ability, who has shown by words and actions that a desire for entry to the teaching profession has been motivated by the highest principles. The writer feels that, provided certain economic difficulties could be overcome, here, in the Emergency Scheme, is a type of candidate who, in order to fulfil the urgent needs of a post-war world has shown himself capable of training, and of playing an important part in the educational system of the country, and who could, under normal circumstances, continue to find a place in the candidature for this profession.

Acknowledgement is due to the Principals of the four training colleges concerned for their co-operation in this investigation, and to the students for their willingness and thoroughness in giving responses to the Questionnaire.

THE VISUAL PRESENTATION OF FACTUAL DATA.[1]

By M. D. VERNON
(*From the Psychological Laboratory, University of Reading*).

I.—*Introduction.* II.—*Method of Experiment A.* III.—*Results of Experiment A*: (1) *Answers to factual questions*; (2) *Answers to general questions.* IV.—*Method of Experiment B.* V.—*General conclusions.* VI.—*Summary.*

I.—INTRODUCTION.

It was shown in a previous paper by the present writer[2] that adult subjects to whom numerical data were presented in graphical form often had considerable difficulty in understanding and remembering such data. Their ability varied with their education and intelligence. The less educated frequently became quite confused as to what the graphical data were intended to convey. And even the better educated, who could grasp and remember the content of the graphs singly, showed little skill in generalizing from a series of graphs to produce a continuous statement or argument embracing the information given in the whole series.

It was decided to extend the enquiry to discover how far the same type of difficulty was found in grammar school pupils; and also to find the result of presenting a written text together with graphical material similar to that used in the earlier experiments.

II.—METHOD OF EXPERIMENT A.

The material used consisted of data relating to vital statistics—on population, mortality, etc. Material of this type had been found to give good results in the previous enquiry. It was thought probable that it would be unfamiliar to the majority of those tested, which indeed proved to be the case. The material was divided into three series, each of five presentations, which were as follows:

Series I
(1) Change in size of population of various countries.
(2) Variation in the birth-rate and death-rate of this country since 1851.
(3) Variation in the net reproduction rate since 1851.
(4) Variation of the birth-rate in different social classes.
(5) Variation in the size of the age-groups since 1841.

Series II
(1) Death-rates of men and women at different ages.
(2) Death-rates from various diseases in the different social classes.
(3) Relation between death-rate and wages.
(4) Change of death-rate from tuberculosis and rheumatic fever, and of weekly wages, since 1875.
(5) Death-rate from tuberculosis of town and country dwellers.

Series III
(1) Decrease of infant mortality in various countries since 1871.
(2) Variation of infant mortality from various diseases in different social classes.
(3) Infant mortality and nutrition in various countries.
(4) Relation of diet to incidence of various complaints.
(5) Relation of social class to various diseases.

Each of these presentations was prepared in three forms: a pictorial chart or pictogram, a graph and a table of figures. By having three groups of subjects, it was possible to rotate the series and the forms of presentation so that each series was seen by an equal number of subjects in each of the three forms. Each presentation was headed by a statement of its contents. As far as possible, the technical terms such as birth-rate, net reproduction rate, etc., were explained in familiar words.

---

[1] The content of this paper covers the same ground as that of a paper presented to the British Psychological Society at the Extended General Meeting held in Reading in March, 1950.

[2] In the *British Journal of Psychology*, 1946, XXXVI, 145-58.

The presentations were shown individually to each subject, each for a period of two minutes. The subject was required to study the presentation carefully; and after it had been removed, he was asked orally three or four questions relating to the specific facts shown in it. At the end of each series of five presentations, he was asked one or two more general questions designed to cover the content of the five presentations. These general questions were as follow:

Series I.	" What changes do you think are happening to the population of this country, as shown in these charts (or graphs, or tables of figures) ? "
Series II.	" Who, according to these charts (graphs, tables of figures), are the most healthy people in the country ? What social conditions seem to be the most important in producing health ? "
Series III.	" Do you consider that the health of children in this country is satisfactory, as shown in these charts (graphs, tables of figures) ? How could it be improved ? "

The time allowed for answering the questions was not limited. Answers were recorded by the experimenter. At the close of the experiment, the subject was asked questions relating to ease or difficulty of understanding the data, preference for one form of presentation or another, etc. The subjects were twenty-four boys and fifteen girls aged 15-18 from the sixth forms of grammar schools.

III.—RESULTS OF EXPERIMENT A.

(1) *Answers to Factual Questions.*—One of the objects of this investigation was to discover whether the form of presentation affected the ease of understanding and accuracy of remembering the data presented. It is clear from Table I that, on the average, there was no significant difference between the three forms of presentation, pictorial chart, graph or table of figures, as regards the accuracy with which the questions were answered. But in certain particular cases one of the three forms showed a distinct advantage over the others. Thus Fig. 1 (a) and (b) illustrate a case in which 73 per cent. of the questions were answered correctly by those shown the graph, and only 55 per cent. by those shown the chart. However, the table of figures gave a slightly better result than either, producing 78 per cent. of correct answers. On the other hand, with Fig. 2 (a) and (b) the chart produced more correct answers (80 per cent.) than did the graph (58 per cent.)

TABLE I

	Average per cent. correct answers		
	Boys		Girls
Charts	71·5		69·3
Graphs	77·9	t=1·16	71·0
Tables of figures	78·1	P=0·26	69·3

Figure I (a)

POPULATION CHANGES IN VARIOUS COUNTRIES

This Chart shows the sizes of the populations in various countries in the years 1920 and 1939.

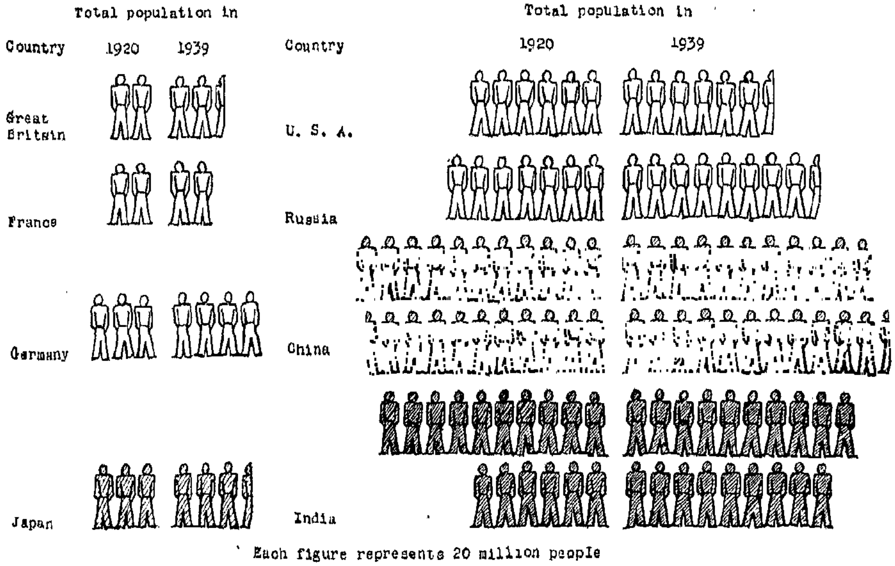


Figure I (b)

POPULATION CHANGES IN VARIOUS COUNTRIES

This graph shows the sizes of the populations in various countries in the years 1920 and 1939.

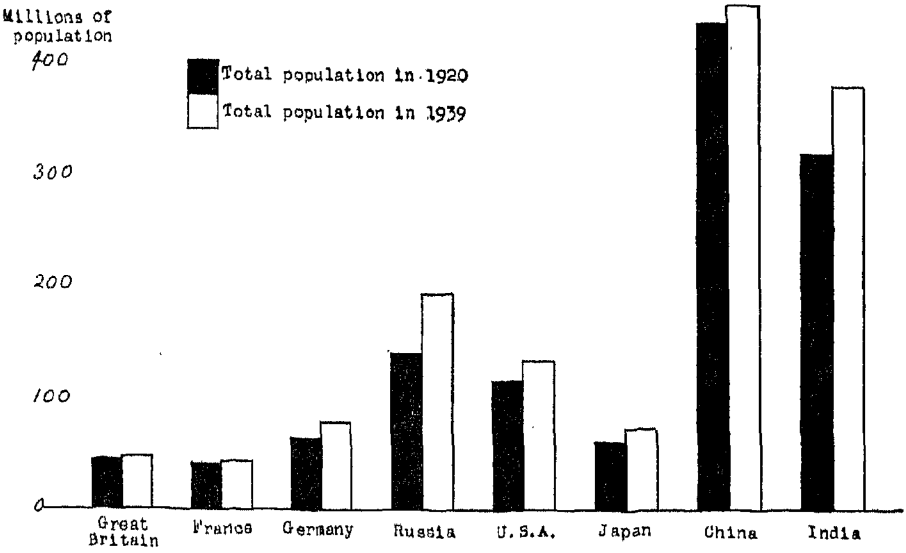
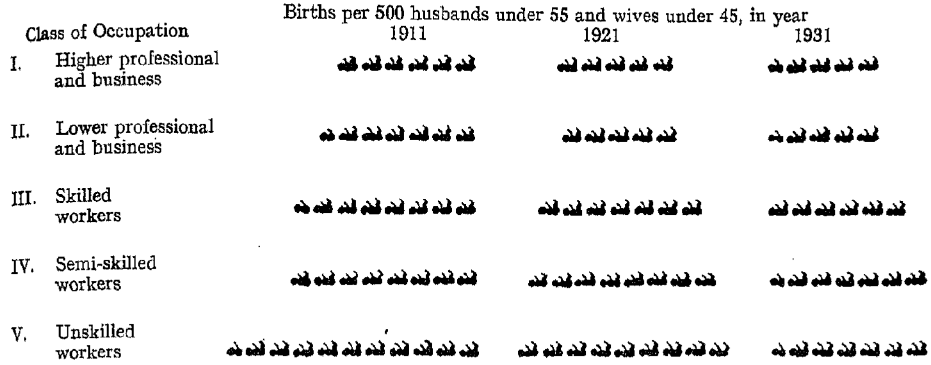


Figure 2 (a)

THE VARIATION OF THE BIRTH-RATE IN DIFFERENT SOCIAL CLASSES IN ENGLAND AND WALES

This chart shows the birth-rate in five different classes of occupation in England and Wales: the higher professional and business, the lower professional and business, skilled workers, semi-skilled workers and unskilled workers. These correspond to the five main social classes of the population, in descending order. The birth-rate is shown for 1911, 1921 and 1931.

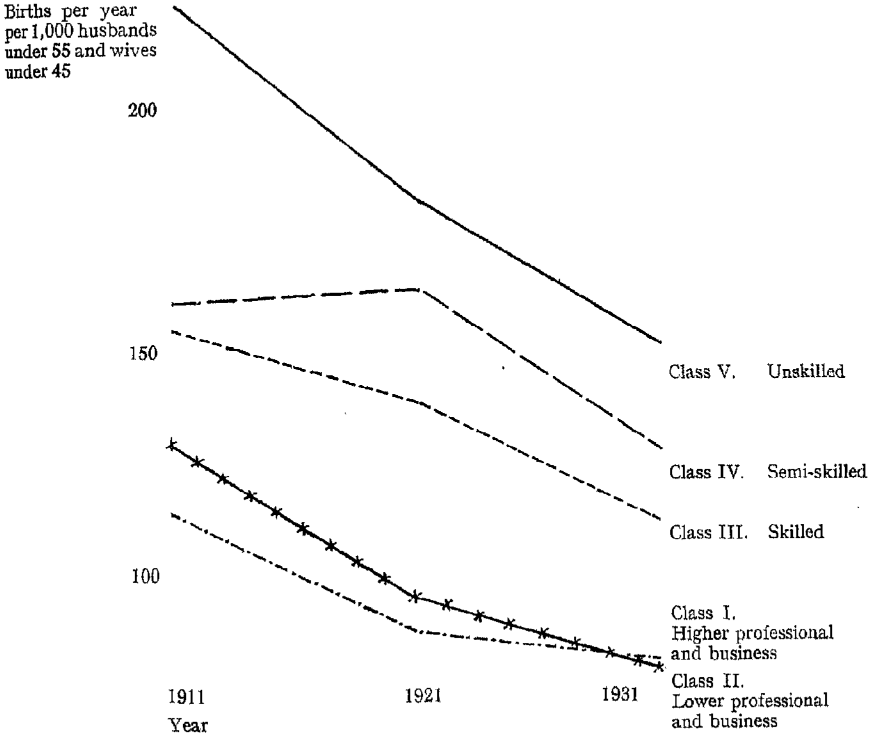


Each baby represents 1 birth.

Figure 2 (b)

THE VARIATION OF THE BIRTH-RATE IN DIFFERENT SOCIAL CLASSES IN ENGLAND AND WALES

This graph shows the birth-rate, the number of births among 1,000 people of the population, in five different classes of occupation in England and Wales: the higher professional and business, the lower professional and business, skilled workers, semi-skilled workers, and unskilled workers. These correspond to the five main social classes of the population, in descending order. The birth-rate is shown for 1911, 1921 and 1931.

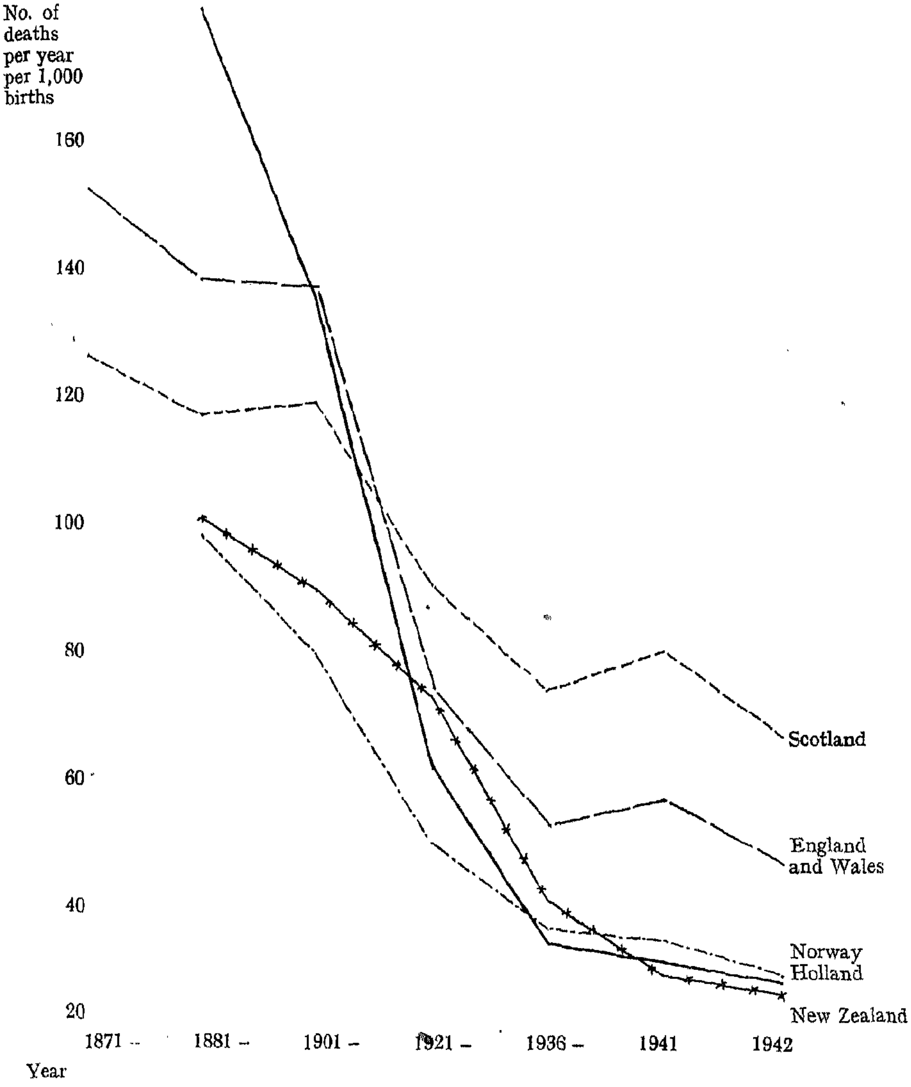


It is difficult to lay down any exact principles to determine whether in any particular case a graph or a chart will give better results, and in general it is better to make an empirical test on this point. However, the graph gives a more delicate and accurate representation of small differences of value, especially in cases where first one frequency is greater and then another—that is to say, where the lines of the graph cross each other. Such a graph is shown in Fig. 3. But the

Figure 3

INFANT MORTALITY IN VARIOUS COUNTRIES

This graph shows the infant mortality, that is, the number of deaths of children under 2 years, per year, for every 1,000 children born, in various countries, from 1871 to 1942.



finer points of detail are apt to be ignored by the subjects no matter how they are presented. So also are differences of scale between different frequencies. The charts can give a general impression rapidly and strikingly, especially when they are comparatively simple, and the figures are not crowded together (compare for instance Fig. 1 (a) and 2 (a)). Whenever it is important to remember exact numerical values, tables of figures may produce a better result than either graphs or charts. In all cases, the simpler the presentation, the more likely it is to be readily understood and remembered.

When questioned as to whether they preferred graphs or charts or tables of figures, the boys and girls on the whole liked the graphs best, the charts next best, and the tables of figures least. But Table II shows that there was a good deal of variation in opinion, although the difference between the three forms was significant. It is possible that this preference for the graphs may have been due in part to their greater familiarity. But also quite a number of subjects recognized that considerable labour was imposed by counting up the number of symbols in the charts to calculate the frequency values.

TABLE II

	Number of times		
	Preferred 1st	Preferred 2nd	Preferred 3rd
Graphs	14	6	11
Charts	9	10	12
Tables of figures	7	12	11
Total	30	28	34

$X^2 = 4{\cdot}66$, $P = 0{\cdot}03$

However, the ease and difficulty of understanding and remembering the information presented depended to a much greater extent on the intrinsic difficulty of the content. In general, the data on population, birth and death-rates, etc., presented in Series I were considerably harder to understand than those presented in Series II and III. The difference of accuracy in answering between these three series was significant, as is shown in Table III. Again, particular presentations were especially hard to understand. That shown in Fig. 4 (see p. 183), giving the changes in the age group, was especially hard, and with all three forms of presentation only about 40 per cent. of the answers were correct.

TABLE III[1]

	Average per cent. correct answers given by	
	Boys	Girls
Series I	62·1	55·9
Series II	80·2	73·8
Series III	75·4	73·9
	t, I and III, $=3{\cdot}92$ $P < 0{\cdot}01$	t, I and II, $=3{\cdot}53$ $P < 0{\cdot}01$

---

[1] No significance can be attached to differences shown in these tables between boys and girls, since there was a slight difference of procedure between them. The girls saw each set at a separate sitting, the boys all three sets at one sitting.

Nevertheless, on the whole the specific factual questions were quite well answered ; the average percentage of answers correct was about 70. One might conclude from these results that the subjects had gained a reasonably good understanding of these data, or at least of the data shown in Series II and III. Indeed, one might term the results really good when one considers the rather brief time given the subjects to study the data. But if the answers to the general questions are now examined, the conclusions are less satisfactory.

(2) *Answers to General Questions.*—In Section II were given the general questions that were asked at the termination of each series of five presentations ; these questions were intended to evoke conclusions based on the whole series. Also some of the questions asked on specific presentations required not so much a reproduction of information as some inference from this information. In both cases, the results were poor. Thus the question on Series I (2), " Do you think that the number of people in the population will increase or decrease in future ? " produced considerable hesitation and doubt. This appeared to be due partly to a failure to realize that population must increase as long as the birth-rate exceeds the death rate, even when the birth-rate is falling ; and partly to a confusion caused by the reversal of the trend of the birth-rate in 1941. But these failures of understanding and their results were shown up more clearly by the inferential question just quoted than by the preceding factual questions.

It was not possible to score numerically the answers to the general questions. They were graded A, B, C and D, from " full and comprehensive logical statement " to " completely incoherent or incorrect statement." The number of questions falling into the various grades are shown in Table IV. None of the answers was worthy of an A grade ; and as many as a quarter were graded D. Again, the results of Series I, the most difficult series, were the worst. Forty-six per cent. of the answers were graded D, and included statements such as the following (in answer to the question, " What changes do you think are happening to the population of this country ? ") :

TABLE IV

	Boys			Girls			Both		
Grade	B	C	D	B	C	D	B	C	D
Series I	1	9	14	3	8	4	4	17	18
Series II	4	13	7	9	5	1	13	18	8
Series III	10	13	1	8	6	1	18	19	2
Total	15	35	22	20	19	6	35	54	28

$X^2=22{\cdot}73$, $P<0{\cdot}01$

> " Mostly population dropping quite considerably " ; " Going down ; population of older and younger people going up " ; " People having less children ; getting older and also decrease at fifteen " ; " Going up ; getting younger ; net reproduction rate going up."

It seems fairly clear, then, that accuracy in answering specific factual questions does not guarantee a general understanding of the data sufficient to produce a comprehensive statement covering the important facts and relating them coherently together. If the inter-relationship between the separate presentation of a series is, on the whole, obvious and easy to grasp, as with Series II and III, then a fairly reasonable set of conclusions may be made.

Here is an example of such a statement of conclusions, graded B, in answer to the questions: (1) "Do you consider the health of children in this country is satisfactory?"; (2) "How could it be improved?"

> (1) "It is very uneven. For the higher classes it is satisfactory, for the lower unsatisfactory. It is probably better now there is rationing." (2) "By a more scientifically balanced diet; better housing conditions; more people living in the country and in rural conditions; a more efficient hospital service to cure non-fatal diseases; state medical research into incurable diseases; and education comes in too."

Here we see that a list of the conditions shown in the presentations is given, together with a few original suggestions. But the separate items are not related in any continuous and coherent argument, and the suggestions are thrown out in a vague way without any attempt to explain how and in what circumstances they would benefit children's health.

Whenever there was considerable difficulty in relating the data together, and particularly if they showed signs of conflicting with one another, as in Series I, then the task was beyond the ability of these grammar school boys and girls, at least under conditions such as those of this experiment. They picked out a few points and ignored or made contradictory statements about others. Or they produced vague and often incorrect generalizations such as those quoted above. Thus it appears that they required either more assistance or further training in gaining a generalized and coherent understanding of the whole argument presented by the data. It seemed possible that they might find the task easier if they were given a written text to read as well as the separate and discrete presentations; and an attempt to investigate this possibility was made in Experiment B.

IV.—METHOD OF EXPERIMENT B.

The content of the material used in this experiment was almost the same as that of Experiment A. Fifteen charts were employed, twelve of which were the same as those of Experiment A. The following charts were new:

Series I.	(3)	Change in number of children per married couple in manual and non-manual occupations from 1900-1929.
Series II.	(1)	Relationship of adult mortality to overcrowding.
Series III.	(5)	Relationship of infant mortality to overcrowding and to climatic conditions.

The fifteen charts were grouped in three series as before, and a written text prepared corresponding to each series, which contained much the same information as that given in the charts, but with certain differences (omissions and contradictions) on points of detail.

The sixteen new subjects (sixth form grammar school girls, aged 15-17½) were divided into three groups, and tested individually. The first read the text of Series I first for five minutes, and then studied the charts separately for five minutes; read the text of Series II and studied the charts simultaneously for ten minutes; and studied the charts of Series III for five minutes, afterwards reading the text for five minutes. The other two groups of subjects were rotated in such a way that each series was seen in each of the possible arrangements by an approximately equal number of subjects. After each series, questions were asked and answered orally. These were slightly more general than the specific questions of Experiment A, but less general than the general questions. This procedure of reproducing all the information at the close of the series, necessitated by the use of the continuous written text, may have imposed more strain on the subjects' memory that did the procedure of Experiemnt A. On the other hand, it gave the subject the advantage of considering the various items in context with one another.

V.—Results of experiment B.

It might be considered that the procedure in which charts and text were studied simultaneously would have a distinct advantage over the procedures in which they were studied successively. However, this was not the case; the average per cent. of correct answers was 51·8, 51·9 and 53·1 for the three procedures (all differences being insignificant). But it was observed that many of the subjects, even when given text and charts together, read through the text first, then laid it aside and studied the charts in isolation afterwards. Thus they had not apparently learnt how to check the text by means of the illustrative data presented in the charts. But a small group of university students tested subsequently did carry out such a check, and the superiority of their results to those of the school girls was in part due to this. (The students' average score was 69 per cent. of answers correct, as against 50 per cent. for the school girls.)

Once more the material of Series I was found to be distinctly harder than that of Series II and III. The scores of correct answers are shown in Table V.

TABLE V

	Per cent. correct answers for		
	Series I	Series II	Series III
	47·7	56·6	52·6
	I and II	I and III	II and III
Difference	8·9	4·9	4·0
t	2·34	1·13	1·06
P	0·03	0·27	0·30

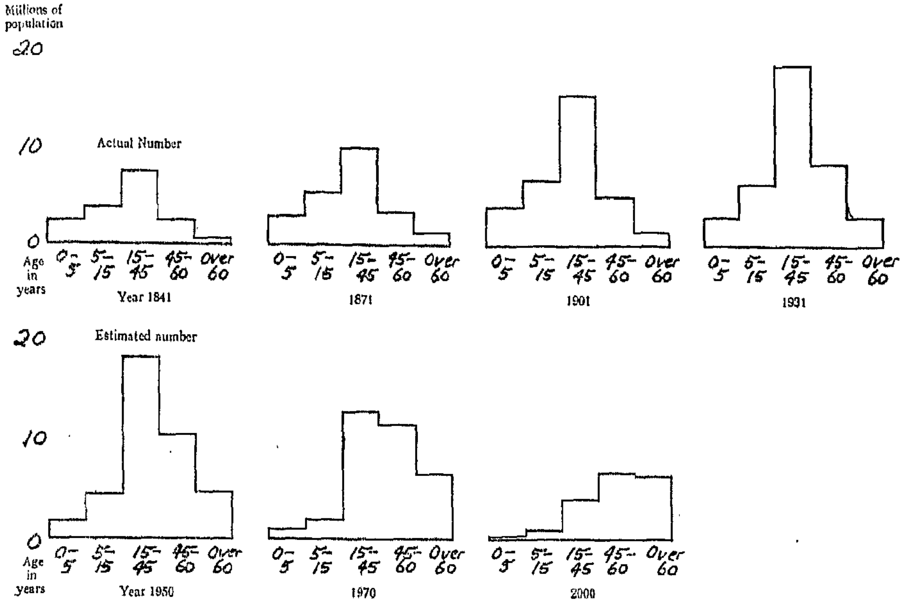
The difference of scores between Series I and III is not fully significant. But when the actual answers were studied, it was found that the mistakes in Series II and III were mainly caused by the forgetting of certain specific facts, and only rarely by a total misunderstanding of the material. The answers to the questions on Series I often showed a real confusion about the information given, and even a contradiction between one answer and another. As was mentioned previously, there was a frequent misunderstanding of the fact that the population has risen although the birth-rate has fallen because the latter has continued to exceed the death rate. Further difficulties were caused by the reversal of the trend of the birth-rate during the war; and by the fact that some of the charts (for instance, that on ageing, the data of which are shown in the graph in Fig. 4) were based on pre-war figures which did not take this reversal into account. The result was that in several cases the reversal was noted by the subject in some of her answers and not in others.

A further source of confusion was deliberately introduced into the experiment by the lack of complete correspondence between the text and the charts. There were a certain number of actual disagreements over minor details; for instance, the text stated that the population of France was decreasing, while the chart showed it to be very slightly increasing. In other cases, small factual details given in the charts were omitted from the text, or *vice versa*. Also, certain explanatory statements given in the text (such as that " the higher infant mortality among the poorer classes from diarrhœa and enteritis is probably due to unsuitable feeding ") were of necessity omitted from the charts.

Figure 4

THE CHANGE IN THE SIZE OF AGE GROUPS IN ENGLAND AND WALES

This graph shows the number of people of different ages there were in the total population of England and Wales in 1841, 1871, 1901 and 1931; and it also shows the number there would be in 1950, 1970 and 2000 if the birth rate and death rate went on falling as they did before the war.



It was interesting to find that only three of the sixteen subjects noted any actual contradiction between charts and text, and only one of them could remember a particular instance. Nevertheless, it seemed that these contradictions increased the subjects' confusion. Moreover, the contradictory items, and also details appearing in the text or charts only, were less well remembered than items for which there was a corroboration between text and charts. There was no consistent tendency to remember isolated details given in the charts rather than those given in the text alone, or *vice versa*. But some of the explanatory statements made in the text seemed to be definitely valuable in providing a meaningful context and giving a reason for the occurrence of the facts presented.

Such inferential and explanatory propositions, which can be included in a written text, but not in the graphical presentation of the data, are probably valuable to individuals who lack the necessary background of knowledge and skill in reasoning to provide explanations for themselves. This lack was demonstrated in this experiment in the answers given to questions which demanded inferences, such as: "Do you think that the population is likely to increase or decrease in future? What makes you think this?" The following were some of the answers given:

"Because of more stable conditions, will increase." "Depends on the state of the country—people don't have enough money to have a lot of children." "Increase, because there's no war at the moment." "Won't increase, because it's so hard to maintain a family." "Increase quite likely, because there are so many modern conveniences and helps." "Decrease when people realize it's harder to get children educated properly." "Likely to increase because married couples want to bring up children to help them in their homes." "Increase while there is still peace, because while people have something worth living for, they have children more easily."

It is clear from these answers that the subjects were unacquainted with such facts as that the birth-rate is liable to fall with a rise in the standard of living, and even with such stable conditions of livelihood as exist in Sweden. Two of the charts showed that in the past the higher social classes had a lower birth-rate than the lower social classes. But the subjects were unable to reason from this to the unlikelihood of better social conditions producing a growing population, in view of the nearly complete constancy of the death-rate.

VI.—General conclusions.

Once more, then, we reach the conclusion that the essential factor in understanding and remembering such material as this is the production of a continuous and coherent argument, connecting together the isolated factual data. The highly educated and intelligent individual may be able to go some way towards providing such an argument for himself. Some of the university students tested showed signs of doing this. But, in general, the evidence was that it was beyond the powers of individuals with less intelligence and education than that of the good university student. Hence, it is necessary that such an argument be presented to the individual and fully grasped by him if he is to gain a real understanding of the information presented and the conclusions to be drawn from it.

We cannot make any very definite statements on the basis of these particular experiments and their results as to the best method of presenting such an argument. It seems reasonable to suppose that it should be verbally presented, since only in words can reasons and relationships be directly stated. Also, it seemed helpful to include certain explanatory inferential statements in the argument. These may, of course, be no more than suppositions, as for instance in the statement about the cause of diarrhœa and enteritis which was quoted above. In a more detailed treatment of such a topic, it would be desirable to give more evidence for such suppositions. But they should not be omitted altogether; otherwise the subject, ignorant of any apparent reason for the occurrence of the facts stated, is liable to forget them, or to misunderstand the general course of the argument, or to invent much less reasonable suppositions for himself.

As to the value of presenting diagrams or tabular material with the written text, these experiments again offer no definite evidence. It is hoped to investigate the problem more fully in future. But it seems justifiable to conclude that if illustrative material is introduced, it must be simple and readily comprehensible in itself, and must not offer any point of disagreement with the written text. It often happens that details are presented in diagrams which are not mentioned in the text, with the object of giving further factual evidence for the textual statements. Again, the well-educated individual can assimilate such additional facts to the main argument. But is possible that anyone who is less highly skilled may find them in a source of confusion because he cannot integrate them within the main argument of the text.

VI.—Summary.

1.—Groups of boys and girls from the sixth forms of grammar schools were shown sets of data on vital statistics: A—In each of three forms, as pictorial charts, graphs, and tables of figures; B—As charts accompanied by a written text. The subjects were required to study these data and answer questions about them afterwards.

2.—There was little difference on the average in the results produced by each of the three forms in A. But one form might have the advantage in a particular case where it presented the data more simply and clearly than did the other two forms.

3.—The main factor determining the accuracy of remembering of particular presentations was the intrinsic clarity and comprehensibility of the content of the presentation.

4.—This factor operated even more strongly in determining the subject's ability to draw general and logically related conclusions from a series of presentations. None of the school children showed great ability in producing a coherent statement of general conclusions.

5.—When a written text was studied as well as the graphical presentations, the same difficulties appeared, and were enhanced whenever there was any contradiction between the statements of the text, or between the text and the graphically presented data. Little ability was shown in checking the text against the charts.

6.—It was concluded that for a thorough assimilation of information of this kind, it is necessary that a coherent and logical statement of the argument should be presented to the subject, including explanatory statements to integrate the discrete facts together.

THE EFFECTS OF PUPIL-MIGRANCY ON ABILITY AND ACHIEVEMENT IN RURAL AND BURGH SCHOOLS

By A. C. PATTERSON
(*English Department, Leith Academy*).

I.—*Purpose of study.* II.—*The data.* III.—*Preliminary analysis.* IV.—*The candidates: (a) Age; (b) Migrancy.* V.—*Achievement in English and in Arithmetic.* VI.—*Teachers' Estimates.* VII.—*Conclusions.*

I.—PURPOSE OF STUDY.

THIS enquiry attempts to examine certain aspects of conditions obtaining in small rural schools and in larger burgh schools of the same area of Scotland. The basis of the analysis is certain pre-war Control Examination data. This Control Examination, or Qualifying Examination as it is sometimes called, is held at the end of the primary course; in this area it must be taken by all pupils aged 12+, irrespective of class placement. It takes the form of standardized tests of the Moray House series, each pupil being tested in Intelligence, English, Arithmetic. On the results of these tests pupils are awarded places in the senior or junior secondary schools. By analysing tabulations of these examination results, the writer hopes to place on record a comparison of rural with burgh schools of the same area and time. In these rural schools, which are all one- and two-teacher schools, each teacher is responsible for the supervision and instruction of more than one class or age-group at the same time; in the one-teacher schools there may be taught in one room a complex of classes, ranging from infants to post-control. That is, one teacher is required to teach simultaneously six, seven or eight classes. In two-teacher schools the number of classes decreases to three or four to compensate for the increased numbers in each class. The set curriculum is embodied in an official county scheme of work leading to the annual Control Examination, the requirements of which are identical for all schools in the area concerned, and is expected to be completed by the day of the examination by all examinees, whether taught in a rural classroom under the conditions outlined or in a burgh school where each teacher is responsible for the work of one class only. Can a rural teacher in the same fixed time cover the common curriculum as efficiently as the burgh teacher whose efforts are not distracted by the needs of other classes in his or her charge? In short, are the two types of school equally efficient in the preparation of candidates for this specific examination held at the end of the primary course? It is hoped, by analysing the available data, to attempt to answer this question and to assess whatever factors may be found to operate favourably or otherwise for the rural teachers.

II.—THE DATA.

The material used was generously provided by the authority concerned. It consists of the pertinent extracts from the information presented to the County's Examination Board in 1938 and 1939 for the assessment of Control Examination candidates. Details for the purpose of this enquiry are grouped separately for (1) One-teacher schools; (2) two-teacher schools; (3) burgh schools.

For both years and for the three types of school the schedules show for each pupil presented: (1) Age in years and months; (2) numbers of changes of school; (3) English record; (4) Arithmetic record; (5) I.Q.; (6) E.Q.; (7) A.Q.

In addition, the 1939 schedule includes a conversion of the raw scores in English and in Arithmetic to a county scale mark.

Though these facts are shown on the schedules, it is to be noted that the names of schools and pupils were not divulged, schools being listed as A, B, C, etc., and pupils being entirely anonymous.

The number of candidates presented by all three types of school in 1938 and 1939 were 303 and 308, respectively. In 1938 the burgh schools presented 209, and the rural schools[1] 94. For 1939 the numbers of candidates were burgh 212, rural 96.

III.—PRELIMINARY ANALYSIS.

TABLE I

THE MEAN AND THE STANDARD DEVIATION OF I.Q., E.Q., A.Q., FOR 1938, 1939, SHOWN SEPARATELY FOR RURAL AND BURGH SCHOOLS.

	Rural N=190		Burgh N=421	
	Mean	σ	Mean	σ
I.Q.	94·88	13·9	102·08	14·9
E.Q.	97·13	13·9	105·04	15·7
A.Q.	99·92	14·2	105·28	14·0

In each of the three quotients the rural schools are inferior to the burgh, though the distribution is more compact. The difference in I.Q., 7·2 points, is much greater than might be expected.[2] However, examination of these quotients in relation to the I.Q., suggests that the teaching techniques of the rural teacher have a large degree of success in handling this situation. If the I.Q. is taken as standard, E.Q. is +2·2, A.Q. +5·0 for rural schools while those for the burgh schools are +2·9 and +3·2.

Before accepting these gains as valid, two criticisms suggest themselves:

(1) They may be the result of some limitation of the standardisation of the norms of the tests used.

(2) They may be the product of coaching in the tests used, or in similar tests.

The first point is not difficult to dispose of. The particular tests used were variants in the Moray House series. The 1938 tests were standardised on the norms (N=6,879) of an English county partly rural and partly industrial. The 1939 series were based on complete year groups: The Intelligence test norms (N=11,418) were standardised on both English and Scottish pupils, as were also English and Arithmetic (N=9,657).

The second criticism, the possibility of specific coaching, is more difficult to meet. There is some indirect evidence that such coaching took place both in town and country schools, but it is impossible to estimate the extent. Strict precautions were taken to ensure that no unused copies of the test papers were retained in the schools. In 1949, some two months after the examination, about 10 per cent. of the candidates were re-tested, and a general, though irregular increase in quotients was found. This does not suggest that, generally, coaching was excessive. It must, however, be admitted that this need not necessarily be

[1] This title henceforward will denote schools of the one- and two-teacher type.

[2] A. S. MOWAT: *City and Rural Schools*, Scottish Council for Research. University of London Press, 1938, p. 63, Table XXI (b); p. 38, Table IX.

true of ten years earlier, though it probably is. Alternatively, if coaching be admitted as a significant factor in the gains, instead of claiming that rural schools teach English and Arithmetic better, relatively to endowment, it could be legitimately argued that, relatively to the educational attainment of the pupils, rural schools coach less on intelligence tests than do the burgh schools—a somewhat unlikely conclusion. It is pertinent to draw attention here to some of the findings of Hugh McRae in his study, " The Inconsistency of Group Test I.Q.s."[1] He claims that practice and coaching have similar effect in appreciably raising the median of a group, but that one preliminary test does much to offset the effect of previous practice or coaching and that this compensation is more marked in the case of tests incorporating practice material. The Moray House Tests are preceded on the day before the examination by a practice sheet. For the purpose of this enquiry, therefore, it is claimed that though some coaching has been given, the gains are substantially true gains and represent a definite relationship of one type of school to the other.

TABLE II

Rural and Burgh I.Q.s for 1938, 1939, arranged in Class-Intervals by Numbers and by Percentages.

	Rural		Burgh	
Quotients	Nos.	%	Nos.	%
130+	1	·5	5	1·2
120—129	12	6·3	50	11·9
110—119	18	9·5	86	20·5
100—109	30	15·8	95	22·6
90—99	51	26·8	93	22·1
80—89	56	29·5	65	15·2
70—79	22	11·6	27	6·4

The burgh schools have 56·2 per cent. of their candidates with an I.Q. of 100 or better. The corresponding rural percentage is 32·1 ; that is, the rural teacher may expect 24·1 per cent. fewer " good " candidates and 24·2 per cent. more " poor " candidates than his burgh colleague. This is at variance with both Miss Macmeeken[2] and with Professor Godfrey Thomson.[3] To meet the first, it is suggested that the smaller deviation of the rural intelligence is due, not to an evenclustering of rural I.Q.s round the mean, but to an overgreat squeezing of cases into the under 100 range. In the second case, these remote rural schools, unlike those in the Cheviot valleys, have a marked minority with I.Q. 100 or better. In this area, differences in race and social standing most probably account for the difference. Outlying farms in this district rely a good deal on Irish workers and on the short-term service of migrant labour. Neither of these classes is as well-off financially as the artisans and tradesmen of the burghs.

So far it is established : (1) The rural schools of this area in the two years 1938, 1939, were presenting candidates who, as a group, contained a much larger number of " poor " pupils than were presented by the burgh schools ; (2) the teaching methods of the rural schools was able to advance the achievement in English and Arithmetic of these " poor " candidates to a greater degree than were those of the burgh schools.

---

[1] H. McRae : " The Inconsistency of Group Test I.Qs." This *Journal*, Vol. XII : Part I. February, 1942, pp. 62, 63.

[2] A. M. Macmeeken : *The Intelligence of a Representative Group of Scottish Children* ; Scottish Council for Research in Education. University of London Press, 1939, p. 39.

[3] Godfrey H. Thomson : " The Northumberland Mental Tests." *Brit. Jour. of Psych.: General Section*, Vol. XII, Part 3. December, 1921.

Should more detailed study endorse this preliminary survey, then the efficiency of rural school teaching must be ranked as superior to that of the burghs, unless they concentrate more on English and Arithmetic than do the burghs.

IV.—THE CANDIDATES.

(*a*) *Age.*—The mean ages of the rural and burgh candidates are much alike, 147·6 months and 147·0 months, respectively, but the Standard Deviations 10·5 and 8·2, suggest that in the rural schools some special difficulty exists since both types of school have in this area the same entrance dates for infants and the tests are for all pupils not younger than twelve years and not older than thirteen. There is nothing in the data to suggest a systematic presentation of too young candidates and much to suggest that teachers regard migrants as special problems. Pupils who have made a large number of changes of school, particularly those who have moved frequently in the infant and first primary classes, become retarded, reach Primary V (the normal control class) and the examination, later than they should. The mean age for migrants shows this: Rural, 153·1 months; burghs 150·7 months. On the day of the examinations, 1938, 1939, thirty-four of the ninety-four migrant candidates were thirteen years or older.

This movement of pupils from school to school was due to the movement of married farm workers from farm to farm, formerly a common feature of agricultural life in the south of Scotland and North of England. This arises from the short-term contract of such labour, yearly, or half-yearly. When the contract was not renewed the married worker had to vacate the cottage "tied" to his job and had thus to seek employment elsewhere. Where there was a family this entailed its removing from one school to another. This usually involved a considerable delay between leaving one school and enrolling in another. Emergency Orders during the war years prohibited such movement, and control is still exercised by the provisions of the Control of Engagement Order of 1947, but Ministry of Labour exemptions permit an increased freedom of choice of employer and employee. If such pre-war "flittings" in their full degree have permanently disappeared, then the interest of this study will be mainly historical, a record of teaching conditions of the past. If, as is likely enough, control is gradually relinquished, its value will be more practical.

(*b*) *Migrancy.*—An examination of the numbers of changes of school made by these pupils is tabulated in Table III below.

TABLE III

Pupils per Change of School and by Percentages.

Changes of School	Candidates			
	Rural		Burgh	
	Number	%	Number	%
0	108	56·8	321	76·2
1	27	14·2	61	14·5
2	21	11·1	22	5·2
3	16	8·4	13	3·1
4	10	5·3	2	·5
5	2	1·1	1	·25
6	2	1·1	1	·25
7	1	·5	—	—
8	1	·5	—	—
9	1	·5	—	—
10	1	·5	—	—

Considering as a migrant a pupil who has made more than one change of school, rural candidates include 55 migrants, 28·95 per cent. of the 190 presented. Of the burgh 421 candidates 39, or 9·26 per cent., are migrants. If the table is again examined, the problem for the rural schools appears more serious than the preceding percentages indicate. The incidence of pupils making a large number of changes is heavier in the rural schools. If the product of the number of pupils making each change and the changes involved, greater than one change, is divided by the total pupils, the coefficients emphasise this aspect: For rural schools, ·98; for burgh schools, ·24. From the point of view of stability of the roll, 71 per cent. of rural pupils make no more than one change. In the burgh schools this rises to 91 per cent.

This difference is so marked and so fundamental that it is proposed to take it into account in all further estimates of the work of these two types of school. Thus, a re-examination of the I.Q.s presented in Table I is necessary.

TABLE IV

RURAL AND BURGH I.QS.: MIGRANT AND NON-MIGRANT SHOWN SEPARATELY.

Rural		Burgh	
Migrant	Non-migrant	Migrant	Non-migrant
91·8	98·7	98·9	102·4

In both types of school the mean migrant quotient is less than the mean non-migrant: –6·9; –3·5. This difference represents a factor which must be taken into account when estimating the relative efficiency of these schools. In rural areas migrancy is more markedly present and so affects the distribution of age and intelligence that it represents a characteristic, though not peculiar, problem to the teacher in these smaller schools.

V.—ACHIEVEMENT IN ENGLISH AND IN ARITHMETIC.

There are two possible approaches to an attempted estimation, based on the available data, of the efficiency of each of the two types of school. (1) The relation between Intelligence and achievement in English and in Arithmetic. (2) The relation between the Teachers' Records and the County Scale Scores. In both of these approaches due regard must be paid to the factor of migrancy.

TABLE V

MEAN SCORES IN E.Q. AND A.Q. IN RURAL AND BURGH SCHOOLS SHOWN SEPARATELY FOR: (a) MIGRANTS; (b) NON-MIGRANTS; (c) ALL PUPILS.

	Rural			Burgh		
	(a)	(b)	(c)	(a)	(b)	(c)
E.Q.	92·7	101·3	97·1	101·8	105·4	105·0
±I.Q.	+ ·9	+2·6	+2·2	+2·9	+3·0	+2·9
A.Q.	94·6	103·9	99·9	101·1	106·2	105·2
±I.Q.	+2·8	+5·2	5·0	+1·2	+3·8	3·1

These figures suggest that, in both types of school, achievement is appreciably, but not excessively, ahead of endowment. The differences are so small that any further claim would be of very doubtful validity, though attention may be drawn to the low gains in English of the rural migrants. That there is some disturbance of the normal relationship of I.Q. and E.Q., and I.Q. and A.Q., can be illustrated, perhaps validly in view of the small age range, by correlating (Product-moment) these for the rural schools : r I.Q. E.Q. = ·69 ; r I.Q. A.Q. = ·73 for migrants ; ·81 and ·98 for non-migrants. This suggests that in areas where migrancy is likely to become an appreciable factor in the interpretation of such examination results as these we are considering, then some standardisation of Arithmetic and, still more, of English text-books might be advisable. This would serve to give greater continuity to the work of the migrants.

TABLE VI

THE RELATIONSHIP BETWEEN TEACHERS' ESTIMATES AND COUNTY SCALE SCORES SHOWN SEPARATELY, AS CORRELATION COEFFICIENTS FOR (a) MIGRANTS ; (b) NON-MIGRANTS IN RURAL AND BURGH SCHOOLS.

	Rural		Burgh	
	(a)	(b)	(a)	(b)
English	r ·519	·563	·884	·888
	PE ± ·0701	± ·0580	± ·0330	± ·0219
Arithmetic	r ·469	·494	·872	·783
	PE ± ·0701	± ·0616	± ·0219	± ·0185

It seems from the low correlations and high values of PE in the rural column that little reliance can be placed on the estimates of these teachers when measured thus. The explanation probably lies in the small numbers presented each year with the consequent difficulty of consistent assessment. The burgh schools have, apparently, developed a satisfactory technique of grading and evaluation which copes very satisfactorily with the difficulties of estimating the potentialities of the migrant for this particular examination.

VII.—CONCLUSIONS.

The conclusions which are to be drawn from this study are, of course, particularly applicable to the district furnishing the data, but they may have a wider application to areas elsewhere in which migrancy is to be allowed for.

(1) In English and Arithmetic the ratio of performance to ability is greater in the rural schools than in the burgh schools.
(2) This superiority is achieved in spite of a much heavier incidence of pupil-migration.
(3) Rural teachers might be assisted to do even better by :
  (*a*) Information regarding the technique used in burgh schools for assessing the abilities of both migrants and non-migrants.
  (*b*) Further standardising of text-books and schemes of work, particularly in English.

In closing, the writer would gratefully acknowledge encouragement and advice received from Professor Sir Godfrey Thomson of Edinburgh University. His thanks are also due to the Director and Committee of the authority concerned for generosity in permitting access to the examination material.

CRITICAL NOTICE

Human Ability: By C. SPEARMAN and LL. WYNN JONES. (London: MacMillan and Co., 1950, pp. viii+198, 16s. net.)

Professor Spearman's last work has been eagerly awaited. The greater part was written during the war in frequent consultation with Professor Wynn Jones. After the death of the senior author, Professor Jones completed the manuscript, and prepared it for the press. According to the original plan, the book was to be " a continuation of *The Abilities of Man* "; and the writers claim that " the psychological picture depicted in the older book still remains essentially the same in the new." Much of the volume, therefore, consists of a re-statement of the well known ' two factor theory,' interspersed with detailed references to recent work supporting its various implications.

Several chapters, however, are entirely new. These are mainly concerned with a critical examination of the evidence for supplementary or non-general factors, particularly those which are commonly called ' group factors.' Based as they are largely on work among school children, the views put forward on the existence and nature of specialized abilities –verbal ability, mechanical ability, and the like—should be of particular interest to the educationist. The discussion of statistical methods, which is couched in simple and non-technical terms, will appeal chiefly to those teachers who may be contemplating experimental researches of their own.

The book starts with an introductory chapter reviewing ' pre-scientific notions ' on the problem of mental abilities. As the writers point out, during the closing years of the nineteenth century, a long and inconclusive controversy was waged between the adherents of the traditional theory, which held that the mind included two or three dozen ' intellectual faculties,' and the opponents of the faculty hypothesis, who inclined more and more to the view that there was only one essential cognitive function. Binet is cited as a belated representative of the faculty view, and Spencer as an early advocate of the unitarian view—the champion of a single central faculty which he named intelligence.

However, among the " new psychological monists," as the following chapters point out, very different theories were evolved about the actual nature of this fundamental process. Wundt, Ward, and their respective followers insisted that it was a synthetic activity, broadly identifiable with ' attention ' or ' apperception ' ; Bain, Sully and Spearman himself preferred an analytic interpretation, and emphasized the supreme importance of ' discrimination ' as the most elementary form of cognitive capacity. An objective procedure for resolving these conflicting views at length " arrived " (as Spearman reminds us) " with the genius of Galton, in the form of the statistical device designated correlation." If general ability or intelligence was identical with discrimination, then the correlation between (*a*) assessments for intelligence and (*b*) capacity for sensory discrimination should be positive and, indeed, practically perfect, provided both variables could be " duly corrected for unreliability or error."

Galton himself, however, believed the processes involved to be far more complex than that. In his *Study of Hereditary Genius* he proposed to distinguish two kinds of ability : (i) what he called ' general ability ' and (ii) what he described as ' specific abilities ' or ' special aptitudes,' in addition to (iii) ' errors ' due to ' peculiar or irrelevant conditions '—*three* kinds of converging influences altogether. If, therefore, ' special abilities ' are to be included among the possible factors, we shall only be able to determine ' general intelligence ' after we have eliminated their effects as well as those of the "errors or other irrelevancies." To do this a new technique was needed.

Galton's idea of calculating an index to measure the amount of correlation between *two* variables was extended by his eminent follower and collaborator, Karl Pearson, to the case of *many* variables operating together. Pearson was thus led to introduce, not only the familiar product-moment formula, but also the device of partial correlation. Above all, it was Pearson who first had the idea of reducing a series of correlated measurements, obtained for a number of different traits or tests, to terms of a set of uncorrelated 'index-characters' (or 'factors' as they were subsequently called). In the present volume Galton's contributions are generously acknowledged ; but, as in most of Spearman's later writings, Pearson's early work still remains under a ban. To discover the great debt which educational psychology owes to Pearson, the student of to-day would have to turn back to the older books of Thorndike, Brown and Kelley, which are mostly out of print. However, now that the long and heated controversy between these two protagonists has become a thing of the past, it is to be hoped that Pearson's profound and original suggestions will once more be fully recognized. After all (though younger psychologists often forget it) factor analysis as a statistical method owes far more to Pearson's papers of 1901 than to any later contributions of Spearman.

In his own psychological studies Pearson relied almost exclusively on teachers' assessments. But in America, another pupil of Galton's, J. M. Cattell, administered a series of tests, drawn up in conjunction with Galton himself, to more than 200 students ; and in 1901 there appeared what Spearman describes as " the momentous investigation of Cattell and Wissler." This contained the first published table of correlations between results derived from mental tests, with coefficients calculated by Pearson's procedure. Its object was to " find a means whereby the fundamental elements of general and specific ability could be isolated and valued." But, unfortunately, so Spearman contends, the investigators failed to observe that their results, like those of the 'Galton-Pearson school' generally, were " vitiated by a fatal flaw, that of 'attenuation'."

Spearman himself entered the field in 1904. The article which he then published on 'General Intelligence Objectively Determined and Measured' is familiar to all students of educational psychology ; and in the present volume he still refers us back to it as " the fountain-head " of his theories. His own immediate aim was, if possible, to erect a single cognitive function in place of the multiplicity of cognitive faculties postulated by Galton, Pearson, and their followers ; and, by adopting their correlational technique, to provide objective evidence for the 'unifocal' view which had been so strongly advocated by Spencer and Ward, but hitherto had been left insecurely resting on speculative arguments or appeals to introspection.

Spearman briefly reminds us how his methods differed from those of his immediate predecessors. Rejecting the more complex kind of test employed by Binet, and the more elaborate statistical devices developed by the 'Galton-Pearson school,' he decided to employ three simple tests of sensory discrimination, of a type familiar to the laboratory psychologist, and to correlate the results with independent estimates of intelligence supplied by teachers and with marks obtained in school examinations. On correcting the observed or 'raw' correlations for 'attenuation' (i.e., for the effect of unreliability), he claimed that the final figures turned out to be almost perfect. Neither in the tests nor in the examination results were there any discernible signs of special abilities, such as a 'talent for mathematics' or 'the auditory capacity alleged to underlie aptitude for music,' nor yet of the 'traditional faculties of the older psychologists and educationists.' He therefore concluded that a single general ability was alone responsible for all the observable correlations, and that no room was left for more specialized factors peculiar to small groups of allied activities : " The specific elements of the activity seem in each case to be different from that in any other." Thus, instead of Galton's assumpton of three kinds of ability—general, special and unique—there seemed to be only two. This simplified hypothesis he later termed, in a phrase borrowed from de Sanctis, the " Two Factor Theory." Whatever we may think of the theory to-day, there can be little doubt that its bold simplicity, and the vigour with which it was defended, furnished a most fruitful stimulation to psychological research ; and, coinciding as it did with the publication of Binet's scale of tests, it aroused a widespread interest in the practical measurement of intelligence among school teachers.

In Spearman's view, so he tells us in his latest book, this simple two-factor theory " still holds the key to the situation." Accordingly, in the second and third chapters he sets out a fuller and more up-to-date statement of his fundamental hypothesis, and attempts to reply to criticisms. As Thomson and others pointed out when first the theory was put forward, its novel feature was not (as is so often supposed) that it affirms the existence of a general factor, but rather that it denies a third or intermediate kind of factor, such as those now commonly described as ' group ' or ' bipolar factors.' In replying to these and other criticisms of his views, Spearman now tends to waver a little. In one place he emphasizes " the vanishment of group factors " ; in another he says that his theory " does not require that group factors should be absent *altogether*, but only that they should be so over the limited range required as a jump-off." Nevertheless, though such factors are not in principle excluded, in point of fact, they have, so he declares, proved to be " either small or rare."[1]

Having thus in the opening chapters summarized the general problem, and restated his original theory, Spearman goes on to discuss the methods to be used. This leads him to a new and suggestive discussion of " recent developments in statistical procedure." And here much of what he says is new. He begins with a review of what he describes as ' direct methods '—' simple summation and weighted summation,' ' centroids,' ' principal components,' and ' bifactors.' He puts his finger at once on the essential difference between his own approach and that of other workers in the same field ; as there still seems a good deal of doubt as to how Spearman's methods of analysis diverge from Thurstone's and from mine, it may perhaps be helpful to quote his view of the matter. Burt's method, he says, was " an outstanding attempt at extending the scope of factorizing." " Whereas the older method had dealt with every triad separately, the newer way took the sum or mean of each whole column of correlations together," and the entire table of coefficients was thus treated as a " solid rectangular block." But the attempt, he considers, was bound to fail. Its chief defect is that it entails " filling out the vacant diagonal cells with stop-gaps." That, indeed, was the objection that he originally raised to my procedure when it was first put forward in 1909 ; the " stop-gaps " consisted of the squares of the saturations, and these saturations, he complained, were the very figures to be determined ; hence the calculation rested on a vicious circle. Some years later he himself put forward an ingenious formula which dispenses (at least for the hierarchical case in which alone he is interested) with the insertion of self-correlations ; and this he now takes as furnishing a standard of comparison.

Turning next to Thurstone's centroid equation, " it will," he says, " be noted that this is Burt's formula ; but Thurstone has diverged in the treatment of the diagonal cells . . . Thurstone seems to have ended by preferring whatever [figure] may happen to be highest in the whole column." But once again, whatever device is adopted, the insertion of " stop-gaps " into diagonal cells is distinctly " questionable." Moreover, one " great weakness both of the centroid method and of that called principal components " arises from the fact that " every factor (except the first) produces in about one half of the tests positive values and in the other half negative ones." This, he maintains, is altogether artificial.

To illustrate the differences between these various procedures, he takes a table of Thomson's, and compares the " loadings " furnished by each method with those reached with his own formula. He concludes that " the *absolute* values vary considerably, but not excessively," and that the " *relative* values show themselves to be practically identical." The divergences, he considers, " seem to result from the different fillings of the vacant diagonal places, and these, as we saw, are quite arbitrary."[1]

---

[1] In this passage Spearman remarks that I myself have misquoted him in saying that his theory implied that the supplementary factors were *both* small *and* rare. But the conclusion, which I have cited above from the end of his 1904 article (a conclusion which he himself underlined), is still more sweeping : it implies that they are non-existent. In *The Abilities of Man* we are told that (with the exception of the supplementary general factors, w, o, and p) " *every* s is independent of every other unless they are *closely* similar " (p. 413, my italics) ; this surely implies that group factors, whenever found, are always " narrow " ; and (on p. 241) we are also told that they are " astonishingly rare."

He next turns to 'indirect methods,' such as rotation and the use of oblique axes. Thurstone's use of such procedures is somewhat severely criticized. The arguments are perhaps too briefly stated to be clear or convincing as they stand, but can readily be filled out from what Spearman has said elsewhere. The use of rotation and, indeed, the general concept of 'vectors' of the mind he attributes in the first instance to Garnett. Thurstone's extension of the principle of rotation he considers too subjective. Almost any theory "can hopefully appeal to it for statistical support"; and, if such methods can be made to yield conclusions which are "consistent with almost any theory," then obviously they must "end by offering real assistance to none." As for "the replacement of the usual orthogonal axes by oblique ones" (i.e., of uncorrelated by correlated factors), that, he points out, involves "a very great sacrifice of mathematical simplicity"; and once again lays the way open to producing apparent evidence for almost any conceivable set of abilities.

In regard to results, the most important novelty is perhaps to be found in the chapter on 'the interpretation of *g*,' where Spearman expounds his most recent views on the psychological nature of the general factor which, he maintains, underlies all cognition. Originally, following Bain and others, he thought it could be identified with simple sensory discrimination; later he suggested neural plasticity; a third suggestion was that it was mental energy. After other investigators had shown the superior efficiency of more complex 'relational tests,' such as the test of Analogies, he developed the theory that the general cognitive factor depended essentially on the neogenetic processes, more especially on what he called 'the eduction of relations and correlates.' In the present volume he suggests that even this view was too broad, and finally concludes that *g* is "characterized essentially by the combination of neogenesis *with abstractness*." This further limitation, however, raises a difficulty which is never fully met. The general factor, *g*, is defined as the factor "common to *all* mental tests," by which he means tests of "efficiency rather than of motivation." Now this definition must include tests of simple mental processes, like sensory discrimination or speed of tapping, such as were so freely used in the early days of mental testing. But sensory discrimination and speed of tapping can hardly be said to require an appreciation of 'abstractness,' and in that case it is difficult to understand how 'abstractness' can be put forward as an essential element in *all* the forms of cognition.

The second half of the book is concerned with a re-examination of the various 'group factors' which educational and vocational psychologists have claimed from time to time to have established. In my memorandum to the Consultative Committee of the Board of Education (subsequently reprinted in *The Measurement of Mental Capacity*, pp. 31-48), I had enumerated "at least twelve special capacities or group factors," as sufficiently well established, most of which "seemed curiously reminiscent of the old-fashioned faculties of the phrenologist." The data were derived largely from investigations carried out by research students, most of them teachers in L.C.C. schools and many of them working under the supervision of Spearman himself. Spearman's inferences, however, were very different: "of special abilities," he declared, "there are but the scantiest indications; the modern version of the time-honoured doctrine of the faculties is none the happier for the change of name." The conflict between our interpretations of the same results arose partly from the high level[1] of statistical significance on which Spearman always insisted, and partly from the difference in our methods of analysis. His own method, " used almost throughout the researches," was, so he now reminds us, " the method of partial correlation " ; he calculates the influence of the general factor (g) on the bases of *external* evidence (supplied either by teachers' estimates or by ' reference values ' obtained from trustworthy ' tests of g '), and then ' partials out ' that influence either by the well-known Pearson-Yule formula or by some simplified working procedure derived from it. But that, of course, is not ' factor analysis ' as that phrase is commonly understood : when we speak of ' analysing ' a set of correlated variables into ' factors ' we ordinarily mean a process of determining components by reference to *internal* evidence only, namely, the evidence which is contained in the table of inter-correlations itself.

---

[1] These conclusions are not quite so true of the table as now printed. For instance, for variable 1, where Spearman's equation gives ·496, the method of 'principal components' gives ·662. Professor Wynn Jones was good enough to ask me to check the values as calculated by Spearman in his original draft. I found that his figures for my own procedures required appreciable corrections; even those based on his own saturation formula needed some modification. The effect of the changes has been to increase the differences in the results obtained. It is not quite fair in the present case to treat Spearman's saturation formula as yielding the ideal results. That formula was deduced solely to meet the hierarchical case (in which alone Spearman was at that time interested); hence it is scarcely appropriate for the table he has taken, since that was constructed by Thomson to include three factors, not one.

Spearman seeks to clarify his present views on all these questions by putting forward a more detailed classification of the factors found. First, instead of a single general factor, we are now asked to recognize three[2]—one cognitive (g) and two orectic (p and o). Secondly, it is suggested that we should distinguish two varieties of specific factor—the ' narrow,' which are virtually limited to one test only, and the ' broad,' i.e., those " overlapping specifics " which " cover an appreciably wide ground." Thirdly, in addition to the general and specific factors, he would apparently now admit (at least as possibilities, though not perhaps as facts) two types of intermediate factor—namely, (i) group factors (chiefly cognitive) and (ii) bipolar factors (chiefly orectic). The supplementary factors that he himself has found should, he maintains, be classed as ' broad factors,' and " must be sharply distinguished from what have been called group factors." His reasons are perhaps not altogether convincing. So-called ' group factors,' he explains, " depend, not on any absolute qualities of the abilities at issue," but simply on the tests that happen to have been inserted in the battery.

But surely the same can be said of the " broad specifics." No doubt, the real reason is that, with the aid of this new classification, the results of recent investigators can still be incorporated within the framework of the original two-factor theory.

Elsewhere he explains that " the concept of group factors " may have some use for " statistical manipulation," if not for the " intrinsic description of the specific factors." Accordingly, he proceeds to outline a new method of group factor analysis which he himself has recently devised ; he names it the method of ' excess correlation.' Readers may remember that, in his joint paper with Hart on ' General Ability,' he contrasted his own ' unifocal ' view with the ' multifocal ' theory as accepted by most other writers at that date.[3] On Pearson's multifocal theory (a theory at one time also accepted by Thorndike) there should be low positive correlations between ' unlike traits ' influenced only by a general factor, and high correlations between clusters of ' like traits ' influenced by group factors as well. Spearman accordingly set out a schematic correlation table for three group factors to illustrate the consequences of such a theory ; no such table, he claimed, had ever actually been encountered with the mental measurements as distinct from physical. In the present volume (p. 84) he appends a similar table for three group factors, with numerical instead of symbol coefficients, exhibiting precisely the same arrangement, which he now says has become " extremely common " in work on mental testing. The peculiarity of such tables is " the excess of the intra-group correlations over the inter-group." Accordingly, assuming that the inter-group correlations are all equal, he goes on to develop a formula which will measure the excess due to the several group factors underlying clusters of similar traits. This formula is freely used in the later chapters ; but it is open to the obvious objection that only in quite exceptional cases are the inter-group correlations even approximately equal. What is far more suggestive than the formula itself, however, is the complete change of view that the development of this " new method " implies.

---

[1] In *Abilities of Man*, Spearman took *five* times the probable error (p. 141) ; I had taken three times. In their section on ' statistical policy ' Spearman and Wynn Jones now state the requirement as four times the probable error, and deplore the practice of those statisticians who have ' thrown overboard this high ratio.' What they do not point out is that, with any ratio over one, the presence of a group factor is still more probable than not : to show that the evidence *for* a verbal factor is *not* conclusive is by no means equivalent to showing that the evidence *against* it *is* conclusive.

[2] This is said to have been the conclusion reached in *The Abilities of Man*. But there we are explicitly told that " altogether four " general factors have been established (the fourth is Webb's conative factor, w), and that o is a ' cognitive ' factor.

[3] *Brit. J. Psychol.*, V, 1913, p. 57. Pearson's view referred primarily to bodily measurements rather than to mental. In Spearman's published books and articles, Pearson is practically never mentioned by name ; but (as those who heard the paper given at a meeting of the British Psychological Society will recollect) it was quite clear that he had Pearson's views and criticisms chiefly in mind.

His detailed exposition of the main kinds of special ability opens with a chapter on the ' verbal or linguistic factor.' As he rightly points out, the evidence for such a factor arose first in connection with work on school children, " where verbal activities enter so largely into work in reading, spelling, composition and the like." But this early evidence, so he contends, was vitiated by the fact that it rested on a bipolar factor ; and bipolar factors are inadmissible in the cognitive field.[1] He examines in some detail the more recent contributions by Alexander, Monaghan, McManama, and others ; and finally concludes that " our whole account of the verbal factor ends upon an unsolved discord : both the voices for and against are true so far as they respectively go ; the real question is only as to which and when one theory is more serviceable than the other."

His next chapter deals with ' mechanical ability.' Here he applies his own " excess coefficient " ; and this in his view places the existence of the factor beyond all doubt. But we ought not, he argues, to think of it as essentially a ' spatial factor,' as so many factorists have done. " In regard to *form*," it is a broad factor involving correlate eduction ; it is distinguished from *g*, however (which also involves ' correlate eduction '), " in regard to its material or *content* " : in the case of mechanical ability the content is concrete, whereas in the case of *g* it is abstract.

The remaining " broad qualitative factors " are discussed more briefly. He concludes that there are no " broad " factors underlying sense-perception, motor abilities, mathematical abilities, attention, memory, accuracy, or speed. There may, however, be " narrow " factors entering into (i) speed in certain limited types of task ; (ii) memory or retentivity for certain limited types of material ; (iii) computation in its simplest form, and (iv) the characteristics which Galton summed up by the word ' fluency.' But, in most of these cases, he considers the evidence to be " only permissive, not coercive."

" Very different," he believes, " is the case of what Burt has called bipolar factors : up to now they have attracted scant attention ; but they have quite recently been coming more and more into their proper estate, as research is turning to the other side of human nature, feeling and volition." On the cognitive side, he still maintains that bipolar factors are devoid of meaning : for mental abilities cannot have ' bipolar effects ' any more than they can have ' negative effects.' But here he seems to miss the essential nature of the conception. A bipolar factor which assigns positive saturations to verbal tests and negative saturations to non-verbal tests does not thereby ascribe a negative effect to some particular kind of ability : what it negatives is the verbality, not the effect of the verbality. When we assign a negative measurement to a pupil who is below the average for intelligence, we do not imply that in his case intelligence has a negative effect ; we are merely classifying him as an individual who is below the average for intelligence. And the same holds good with the negative saturations of the bipolar factor. The difficulty doubtless arises because Spearman always thinks of a ' factor ' as a concrete causal agency, not as a principle of classification.

---

[1] Spearman is here referring to the results obtained by myself and my colleagues under the L.C.C. (Miss Carey and Mr. Bradford) in my *Report on the Distribution and Relations of Educational Abilities*, 1917. However, both myself and Miss Carey also used a group factor analysis, which did not entail negative saturations : (cf., for example, *Brit. J. Psych.*, VIII, 1916, p. 180, and refs.). Moreover, a verbal factor was found in a group factor analysis of the original Binet-Simon tests.

One welcome feature that appears again and again throughout the discussion is the frank recognition that factorial and statistical methods alone are not enough. In the past I have myself more than once criticized some of Spearman's interpretations of his factors, because they rested solely on numerical data, and often seemed contradicted by introspective reports.[1] It has always been my firm opinion that no statistical factor can be safely interpreted unless the abstract results of the internal analysis have been confirmed by more concrete evidence, such as comparison with an external criterion based on independent observations, or direct introspective studies of the mental processes involved. Spearman himself now criticizes the work of several of his followers for their neglect of these supplementary procedures. For example, in discussing tests used by Stephenson and others for isolating the ' spatial group factor,' he points out how a little introspection will quickly show that " such tests can readily be performed in two distinct manners : one may be called analytic, in the sense that attention wanders from one element in the figure to another ; the other is comparatively synthetic, in that the figures are mentally grasped in much larger units, sometimes called ' wholes '." Only the former procedure, so he believes, will " *tend to load the processes with G*."[2] And on more than one occasion he insists on the possibility of deciding by introspective methods many of the fundamental questions which factor analysis has so far left unsolved.

The concluding chapter ends on a note of cautious optimism. We are told that the existence of three general factors may now be taken as safely established. All deal with the common factor of mental energy : " *G* represents the amount of mental energy, *P* its inertia, *O* the facility of recuperation after effortful expenditure of energy." Any more precise specification of their nature, Spearman adds, " is still anybody's guess." The guess of McDougall's followers would probably be that they reflect the three main aspects of all mental processes—cognition, affection, and conation. As regards the so-called group-factors, he maintains that further research has, on the whole, tended to consolidate the somewhat negative position taken up in *The Abilities of Man*. However, he seems to admit that the rigid two-factor theory now requires some modification. In the closing paragraph of his preliminary survey (Chapter II) he declares that the " upshot " will require us to accept " (1) a general factor ; (2) an unlimited number of narrow specific factors ; and (3) very few broad group factors." But this, as Spearman's own numbering shows, is no longer a theory of *two* factors, but of *three* factors. The wording, indeed, is almost identical with the formulation of a three factor theory which I myself put forward many years ago[1]; and the resulting picture closely resembles that outlined by Galton towards the close of the previous century.

---

[1] The most important instance occurred in my studies of children's methods of solving relational problems, employed to test intelligence. Here, as a rule, the academic thinker (like the research-students who devise such tests and the students on whom they commonly try them out) reaches his solution by explicitly ' educing ' relations and correlates ; and the university investigator who has had little direct contact with school children is apt to suppose that they too must follow the same procedure. A somewhat similar fallacy was particularly obvious in Spearman's criticisms of Gestalt doctrines, and again in his discussions on the transfer of training, both of which he originally interpreted in terms of an explicit eduction of relations. In factor analysis it is essential to obtain introspective checks from persons of the same age and mental background as those to whom the tests are to be applied. I am tempted to stress this point here, because, if I may judge by the theses which I have had to read as examiner and by the papers submitted by teachers on various problems in educational psychology, nine out of ten investigators seem content to deduce their psychological interpretation of the factors found from a mere inspection of the tests which yield the largest numerical saturations.

[2] I myself remarked on a similar difference of approach which could often be enhanced by changing the form in which the test material was presented. Some children appeared to make explicit attempts at logical analysis and reasoning ; others seemed to rely rather on a kind of schematic apprehension : " They seem to depend on a synthetic activity, comparable to that popularly described as intuition—the activity whereby we implicitly comprehend the meaning or character of a whole, without explicitly analysing it into its component parts of distinctly formulating their relations . . . Thus the pure and abstract thinker is not necessarily to be hailed as the most intelligent " (' Experimental Tests of Higher Mental Processes,' *J. Exp. Ped.*, I, 1911, p. 104).

It may, therefore, be fairly said that Spearman has been gradually moving from the somewhat extreme and paradoxical position that he took up when first defining his original theory to a conception far more in keeping with that of other workers, including some of his former critics. As a result, factor analysis, which at one time furnished problems for such stubborn controversies between its various adherents, is at last beginning to reveal an increasing measure of agreement in respect both of methods and of results.

The whole volume, as much by what it implies as by what it explicitly lays down, provides a comprehensive indication of Spearman's latest views. Here and there its extreme compression tends to obscure rather than to simplify the picture; and, in order to interpret some of the arguments adduced, both reader and reviewer are forced to turn to Spearman's fuller pronouncements in other papers or on earlier occasions. Probably most of these minor obscurities and seeming inconsistencies are due to the fact that the draft of the chapters was left in an unfinished state by the senior author. As it is, the book in its present shape manifestly owes much to the loyal and conscientious care which Professor Wynn Jones has devoted to his task. Students both of education and of psychology, will fell deeply indebted to him for the time and labour he has bestowed, and for the self-sacrificing way in which he has enabled Professor Spearman's final conclusions to be presented in so succinct a form.

CYRIL BURT.

---

[1] E.g., in my memorandum drawn up at the request of the Consultative Committee of the Board of Education and reprinted as Chapter I of their *Report on Psychological Tests of Educable Capacity* (1924), p. 19, or (more briefly) in *Child Study*, IV, 1911, p. 95. Brown and Thomson also predicted the need for some such change. "If hierarchical order departs significantly from perfection, group factors are essential, and no re-writing of equations can eliminate them . . . The theory of two factors must, in this case, be expanded into a theory of three factors" (*Essentials of Mental Measurement*, 1925, Ch. XI, 'The Present Position (1924)').

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES.[1]

A Factorial Analysis of Reading Ability in 10-year-old Primary School Children.

By J. RICHARDSON.

(Abstract of thesis presented as part qualification for the Degree of M.A. in Education, University of Birmingham.)

Reading may best be understood and the nature of its difficulties comprehended by discovering its constituent factors, the mechanisms on which it depends. The present research aimed at doing this by a study of a representative group of 10-year-old primary school children. The investigation was carried out in three primary schools in Smethwick, 260 children completing the battery of twenty-one tests and assessments. These included tests of the cognitive aspects of reading, assessments of attitude, interest and temperament.

Several speed tests were constructed and validated for the battery: (1) A continuous prose reading test consisting of a continuous narrative, graded in difficulty and sub-divided into four sections, questions on the narrative following each section; (2) A silent reading test, the language and material of which were based on the children's daily experience; (3) Visual and auditory discrimination tests of words and word forms; (4) Assessments of experiential background and of attitude to reading.

Existing measures of attitude were found to be unsuitable for use with 10-year-old children. The form of assessment finally used was based on that of Thurstone and Chave (*The Measurement of Attitudes*, 1932) radically adapted, however, for the present purpose. The test and retest reliability (one month's interval) was ·90 for the assessment of experiential background and ·79 for attitude to reading. Validity as ascertained from criteria described in the thesis (teachers' estimates and the assessment of interest by independent observers of the child in interview) was also satisfactory.

Intercorrelations of the twenty-one tests, assessments and ratings were computed for the boys and girls in the group separately. A number of significant differences appeared between the two matrices, boys' correlations being consistently higher than girls'. The matrices were analysed by Thurstone's Centroid Method. In each case it was possible to extract only three factors (on the basis of Ledyard Tucker's and other criteria). Factor loadings were plotted on orthoganal axes and the axes rotated in an attempt to obtain a clear and unique definition of the factors. 64 per cent. of the total test variance was accounted for in the girls' matrix; 68 per cent in the boy's.

Contributions of the factors were:

Factors	I	II	III
Girls' Matrix	47	12	5 per cent.
Boys' Matrix	54	10	5 per cent.

The three factors were identified as (1) general intelligence as estimated by verbal tests; (2) mainly emotional factor of application in and responsiveness to the school situation; (3) very tentatively, a specific factor appearing in the tests concerned with the mechanics of reading.

A further factorial analysis was carried out on the results attained in the test battery, by fifty of the most backward readers, chosen irrespective of sex and irrespective of the cause of their backwardness. Their scores were correlated, analysed by Thurstone's Centroid Method and the loadings of the three factors emerging, rotated. The results were an interesting contrast to those obtained from the analyses of the two main groups. The children whose scores were analysed were all failing in reading; the factors emerging were likely to be suggestive of the general bases of their failure. They were identified as (1) specific verbal ability; (2) non-verbal intelligence; (3) an emotional factor similar to that which appeared in the main analyses.

---

[1] These Outlines must be submitted through the Head of the Department in which the research was carried out.

As a complement to the statistical analysis and filling out the partial picture presented by it, a large number of case studies were made. In addition to the study of individual children, investigation was made into the reasons for differences in reading ability between groups matched for intelligence, but otherwise contrasted in possibly significant respects on their test battery results.

The main results of the enquiry are as follow:

(1) Two factors—intelligence and verbal ability—seem of great importance as general determinants of reading success or failure. The importance of the verbal factor is reinforced by the analyses of the backward readers' group and by the case studies.

(2) Though verbal ability may not itself be analysable in this age group into more primary, general factors, in the individual, it seems that it may be affected by a multiplicity of causes—innate or environmental.

(3) The close association of scholastic success with emotional qualities was re-emphasized. A temperamental factor emerges from the analyses. In the case studies, groups of successful and unsuccessful readers are most clearly differentiated by ratings for general emotional adjustment. In almost every case, school failure was characterized by complete lack of confidence and some emotional disturbance.

(4) Specific ability in auditory and visual perception does not emerge as a general differential cause of reading ability at the age of 10 years; nor, in the group studies, does attitude to reading or home background, although in individual cases, the latter is shown to be of the greatest importance.

The Measurement of Emotional Instability by Means of Objective Tests.

By A. D. B. CLARKE.

(*Abstract of a Thesis for the Ph.D. Degree, University College, London.*)

PROBLEM.—The purpose of the following research was to assess the validity of certain objective tests which have been put forward as methods of assessing what may briefly be designated ' emotional instability.' For working purposes Warren's second definition was accepted. Those who have maintained the existence of a factor of ' general emotionality,' analogous to that of ' general intelligence,' have argued that, in addition to the so-called ' temperamentally defective,' we should recognize the existence of a milder, non-certifiable group consisting of the ' temperamentally unstable ' (just as, in addition to the ' intellectually defective,' we recognize the existence of a non-certifiable group who are ' intellectually dull ').[1] On this hypothesis the emotionally unstable may be regarded as potentially neurotic. Hence, it is of special importance to the Services to have some practicable method for screening out such cases.

TESTS.—The tests employed were all individual tests, and were selected in consultation with Dr. Eysenck as being in his view the most promising for the purpose. The following were eventually adopted.

(1) *Mirror Drawing.*
(2) *Motor Tension* : (a point pressure test).
(3) *Body Sway* : (*a*) spontaneous (' static ataxia ') ; (*b*) induced by suggestion.
(4) *Association Motor Tests* : (*a*) Motor and verbal disturbance with Luria technique ; (*b*) motor and verbal disturbance with electric shock.
(5) *Preference for Colour or Form* : (*a*) Lindberg's ring test ; (*b*) a sorting test.

SUBJECTS.—168 persons were tested, all drawn from the Services. 68 had been diagnosed as neurotic by Army psychiatrists, and the remainder were normal adults used as controls. It was later found that the control group had a slightly higher average intelligence.

RESULTS.—In the tests for emotional instability the difference between the average score obtained by the neurotic and the normal, respectively, was statistically significant in three cases only, namely, for motor disturbance in the Luria test (but not for verbal), for verbal disturbance in the shock test (but not for motor), for spontaneous movement in the body-sway test (but not for suggested movement). The remaining tests showed no significant difference whatever.

Even for these three tests the correlations with the criterion were decidedly low, namely, 0·48 for the Luria test, 0·27 and 0·24 for the other two. Were these tests used for preliminary screening, one case out of every four would be misclassified. On eliminating intelligence by partial correlation, it appeared that the number of misclassifications would be even greater.

CONCLUSION.—None of the tests proposed has a sufficiently high validity coefficient to claim any practical value for purposes of screening or diagnosis. Moreover, being individual tests, and requiring somewhat elaborate apparatus, they would, in their present form, consume as much time, and need as much expert control, as would be required for an ordinary psychiatric interview. However, these defects might no doubt be overcome by further research ; and the detailed findings of the present investigation suggest the more promising lines for improvement.

---

[1] See *Report of Joint Committee on Mental Deficiency* (H.M. Stationery Office, 1929), Pt. I, pp. 8-9, Pt. II, pp. 49f. The concept of the ' emotionally unstable ' appears to have been first introduced in this sense by BURT (*Child Study*, X, 1917, pp. 61-78). Cf. EYSENCK, H. J. : *Dimensions of Personality* (1947), p. 55, footnote 1 ; and WARREN, H. C. : *Dictionary of Psychology* (1934), s.v.

An Investigation of Perceptual Tests as Measures of Temperament.

BY ANN M. GRAVELY.

(*Abstract of a Thesis for the Ph.D. Degree, University College, London.*)

PROBLEM.—The object of the inquiry was to examine the claim that " tests of temperament have reached a stage where they can take their place with the usual tests of cognitive function and special abilities,"[1] and, in particular, that non-verbal (or perceptual) tests may successfully be employed to discriminate between the neurotic and the non-neurotic members of the population and between those suffering from different types of neurotic disorder, such as hysteria and anxiety states. The idea of diagnosing such disorders by means of psychological tests has come much into favour among a small group of psychiatrists and psychologists working in mental hospitals both in this country and in America. On the other hand, the majority of the medical profession, and probably the majority of British psychologists appear to be sceptical of the possiblity of diagnosing mental illnesses by such means.

TESTS.—The majority of the tests employed were suggested by those described by Dr. Eysenck and his fellow-workers at the Maudsley Hospital in their investigations of ' temperamental types.'

(1) *Flicker.*—A measure of the threshold for visual flicker obtained with an electronic ' stroboflash.'

(2) *Dark Adaptation.*—A measure of the threshold of adaptation obtained after 10 minutes in a dark room.

(3) *Autokinetic Illusion.*—A measure of (*a*) the time taken for apparent movements to be noted; (*b*) the amount of apparent movement perceived, and (*c*) the influence of suggestion on the direction of apparent movement.

(4) *Fluctuation of Ambiguous Figures.*—A measure of the rate of fluctuation in the perception of the Necker cube.

(5) *Estimation of Distances.*—A measure of the accuracy of the visual estimation of line-lengths.

(6) *Motor Perseveration.*—Measured by the success with which the subject can write reversed ' S's after writing ordinary S's and *vice versa.*

(7) *Motor Stability.*—A measure of (*a*) the amount of spontaneous movement while standing with eyes closed (so called ' static ataxia ') and (*b*) the amount of bodily sway induced by suggestion (Eysenck's test).

As a check on the perceptual tests, the following were also included.

(8) *Verbal Association Test.*—This was introduced by way of comparing the results of the non-verbal or perceptual tests with those of a standard verbal test : it was used in the ' closed choice ' form (Maller).

(9) *Vocabulary Test of Verbal Intelligence.*

(10) *Questionnaire on Neurotic Symptoms.* (The version adopted was that prepared for use in investigations at the Maudsley Hospital by Dr. Eysenck.)

SUBJECTS.—168 adult men from the Services, aged for the most part between 19 and 31, consisting of (1) 100 normal persons, used as controls, and (2) 68 diagnosed by Army psychiatrists as neurotic and divided into two main ' types,' viz. (*a*) 34 suffering from hysterical conditions and (*b*) 34 suffering from anxiety states.

RESULTS.—(1) *Statistical Discrimination* : (*a*) *Between Neurotic and Normal.* The averages for the two main groups were significantly different in the case of the following tests: (i) dark adaptation; (ii) quickness of autokinetic reaction; (iii) motor perseveration (production score); (iv) static ataxia; (v) word association test; and, of course (vi) the questionnaire on symptoms. No significant differences were found with (i) flicker; (ii) amount of autokinetic reaction; (iii) fluctuations of ambiguous figures; (vi) estimation of distances; (v) body-sway as induced by suggestion.

---

[1] EYSENCK, H. J. : *Dimensions of Personality*, 1947, p. 259. A fuller description of most of the tests here used will be found in that volume. The writer wishes to express her indebtedness to Dr. Eysenck for further explanations and assistance.

(*b*) *Between Hysterical Conditions and Anxiety States.*—No significant differences were found with any of the tests. There is, therefore, in the present experiments no confirmation of the belief that perceptual and other tests of the type here described can be used to discriminate temperamental types, although the contrast between hysteria and anxiety exhibits such differences in an extreme form.

(2) *Correlations.* (*a*) *With External Criterion.* The dichotomy between normal and neurotic subjects was made the basis of a set of biserial correlations for the five tests noted above as yielding a statistically significant discrimination. The questionnaire on symptoms gave a correlation with the criterion of 0·58 (raised to 0·62 when verbal intelligence was partialled out); the other tests gave correlations ranging (after correction) from 0·36 to ·024. Of these the verbal test (word association) gave decidedly the highest correlation (0·34); but was appreciably reduced when intelligence was partialled out. Multiple correlations were calculated by reducing the correlation table to a triangular matrix of semi-partial correlations.[1] It was found that the addition of the two next best tests to the questionnaire on symptoms only raised the correlation by 0·08.

(*b*) *With Internal Criterion.* A factor analysis was carried out by Burt's method of simple summation. A small general factor was found, accounting for 21 per cent. of the total variance. With this non-perseveration (production score) gave the highest saturation and the questionnaire the next highest. The correlation of the general factor with the criterion amounted to 0·64; this seems due chiefly to the presence of the questionnaire.

Conclusion.—Although with several of the tests there are small positive correlations between the results of the test and the subject's neurotic condition, these correlations (with the possible exception of the questionnaire on symptoms) are far too low for the methods to be of any practical value. There is little evidence for the suggestion that non-verbal or perceptual tests might yield better indications of temperamental differences than verbal, at any rate in their present condition.

---

[1] Cf. Lab. Notes on *Multiple Correlations computed by Hierarchical Subtraction*, and Frazer, R. A., Duncan, W. J., and Collar, A. R.: *Elementary Matrices* (1938), pp. 97f.

BOOK REVIEWS.

La Géometrie spontanée de l'enfant: JEAN PIAGET, BARBEL INHELDER and ALINA SZEMINSKA. (Presses Universitaires de France, pp. 514.)

The three psychologists responsible for this research have already made separately or jointly, a vast contribution to child-psychology, including investigations into the development of intelligence, of concepts of time and number, and of moral judgment. The present volume completes a series of three, the first of which reported an extensive research into the development of a child's concept of quantities; the second volume and the present one cover an exhaustive examination of a child's evolving conception of space and spatial relations. Having dealt more generally with the problems involved in the preceding volume, the authors in this investigate more particularly the spontaneous reactions of children to the three dimensional environment which is projected visually in ever-changing relations as the view point moves.

Part I examines this fundamental reaction by, for example, asking a child to model in sand what it has observed through a window (or what was projected in two dimensions on the window). Part II investigates the growth of a concept of distance and of relative lengths. Part III extends the research to the concept of angles and curves, Part IV to that of surface and volume. The conclusions are summarised in Part V, which reviews the stages in the growth of this complex concept of geometrical relations from a child's first reactions to its visual experiences of the world around it.

The book is to be recommended not only because of the value of the contribution it makes, or for the painstaking thoroughness of the investigation it reports in such detail, but also for the lively interest of the verbatim reports of childrens' answers to questions. J.M.

The Educational Development of Children: WILLIAM GLASSEY and EDWARD J. WEEKS. (University of London Press, pp. xiii+246, 8s. 6d.)

The keeping of school cumulative records, of the kind envisaged by the Norwood Committee, was not always begun as a result of a felt need on the part of the teachers. More often, perhaps, it was imposed on them by a psychologist or administrator who did not always pay sufficient attention to the fact that few teachers possess the necessary knowledge or experience to enable them to undertake any but the simplest aspects of the work. To many, the conscientious filling in of these record cards is a formidable task and it is probably no exaggeration to say that very few cards are adequately compiled, except in those parts dealing purely with school progress. Yet Burt, writing on the " Assessment of Personality " (this *Journal* XV, p. 115) gives it as his opinion that the teacher trained in psychological procedure, " may make by far the best judge of the personalities of his pupils."

This book, which deals with all aspects of record making, should help very much in the training of teachers and also aid those practising teachers whose own training left them ill-equipped in the knowledge and techniques required to form a judgment of a child's personality.

Addressed, not to experts but to " the ordinary class teacher," it begins, after an introduction, with a chapter on elementary statistics. Here are dealt with such things as standard deviation, correlation and the properties of the curve of normal distribution. It is simply and lucidly written and is probably the minimum without which the full use of the numerical data of school records is impossible. Such is the aversion to statistics, however, that there are probably many otherwise able teachers who will fail to get beyond this point. This is no criticism of the book's lay-out, for it is difficult to imagine any other place for this chapter.

Following the statistics are chapters on physical qualities, general ability, attainments, special aptitudes, interests, disposition, and the child's environment. There are appendices giving sample record cards, filled in and explained by case histories; there is a useful list of tests, an extensive bibliography and some sets of mathematical tables. A great deal of ground is covered and this very adequately.

It is a book which was very much needed and it is to be hoped that Sir Cyril Burt, who contributes a foreword, is not being too optimistic when he says, " teachers themselves will eagerly welcome this contribution to their everyday work."

L.B.B.

Figure Reasoning Test: JOHN C. DANIELS. (Crosby Lockwood, 1949, 3s. 6d.)

Teachers' Handbook of Test Construction, Marking and Records: JOHN C. DANIELS. (Crosby Lockwood, pp. 79, 4s. 6d.)

The figure Reasoning Test is a non-verbal group test which takes a little over half-an-hour to administer. It has forty-five items and is designed for use with subjects over the age of ten years. Each item consists of two completed series of three diagrams each and a third incomplete series which can be finished by discovering the relationships which exists between the drawings in the completed series. The answer is selected from six possible solutions.

Standardisation was on 4,500 secondary school children and norms are given in the form of standard scores having a mean of 100 and a standard deviation of 15. By extrapolation the norms have been extended beyond the school range to include a rating for adults. There is also a table showing a special classification based on a grammar school population, but this will probably be of little use outside the area where it was standardised as in the population given here 30 per cent. of the pupils score less than 108 and the distribution has been given an artificial normality which would seem to be hardly justified.

The test is easy to administer and very quick and easy to mark. Answers are written on a separate sheet of paper so that the test booklet can be used over and over again. It is a great pity that it has been printed on rather transparent paper which allows one sub-test to interfere with another. The binding, too, could be improved.

In the *Teachers' Handbook of Test Construction, Marking and Records*, Mr. Daniels argues the case for more extensive use of objective tests and of standardised scores in school examinations. The principles involved in the construction of tests are simply explained and examples are given of objective tests in a wide variety of subjects. The simple statistics which are necessary for the conversion of raw scores into standard scores, are handled lucidly at a level which will be understandable by most teachers. There is a useful appendix of work sheets and tables which should reduce computation to the minimum.

Children Learn to Read: DAVID H. RUSSELL. (Ginn, pp. 403, 24s. 6d.)

This book is a comprehensive survey of the most important findings of research and good classroom practice in the teaching of reading. That it refers almost entirely to American work is what one would expect in a book written for American teachers, but in the references British work has not been neglected.

There are three main parts to the book. The first deals with such aspects of the child's reading activity as the sociological, historical, developmental and psychological backgrounds. Some of this will be of only passing interest to British readers because of its purely American implication. The second part deals with the process of learning to read at various age levels and here stresses a fact which receives far too little attention in English schools, namely that " the complete reading act goes beyond recognition and understanding to interpretation and use." The last section examines in some detail such aspects as the development of vocabulary, of word recognition skill, of reading for content and the acquiring of reading tastes. There is also a useful chapter on providing for the individual differences in ability in a typical reading class.

At times the style is a little laboured but it is always understandable at the teacher level. References and bibliography are extensive though, unfortunately, the bulk of the material will not be readily available here. In spite of this, it will be a valuable book for the reading expert.

Activity Methods for Children under Eight: Various authors, edited by CONSTANCE STURMEY. (Evans Bros., pp. 240, 8s. 6d.)

This book consists of a series of articles written by a score of well-known exponents of activity methods. There is some account of the theoretical background of the need for play and activity for the young child, but this is less well done than the bulk of the book which consists largely of lively descriptions of teachers' experiments illustrated by anecdotes and by photographs.

The articles are grouped into five sections. The first describes typical activity programmes in nursery and infant schools in towns and in rural areas. Then there is a section devoted to the change over from formal to informal methods which is followed by a section on religious training. The fourth section deals with the physical needs of the young child and the last with the way in which the separate subjects and skills are dealt with in an activity programme.

Most infant teachers will enjoy this collection and will be stimulated by it. They will also be grateful for the list of books and materials needed in infant schools which is given as an appendix.

The Interpretation of Dreams: SIGMUND FREUD. (Allen and Unwin, pp. 600, 16s. net.)

This is a further reprint of the third English edition, which was a completely revised translation of the eighth German edition. In his foreword (dated 1931) Freud says the book " remains essentially unaltered." The continued demand for the book is a testimony to the author's fame, but the publishers also deserve complimenting on bringing out so large and well-printed an edition at such a reasonable price as present prices go.

Man and his World: *III—The World's Wealth and its Problems*: JAMES MAINWARING. (George Philip and Son, Ltd., pp. 346, vii, 7s. 6d.)

This third volume completes the admirable study Dr. Mainwaring has given of the history of man in relation to his environment. Evidence abounds that the author knows how to apply his psychology to methods not only of teaching facts, but to the stimulation of thought and reasoning about fundamental political and social problems. As before, the illustrations and production are excellent.

OTHER PUBLICATIONS RECEIVED

Fundamentals of General Psychology: (3rd edit.): J. F. DASHIELL. (Pitman, pp. 690, 27s. 6d.)

Adler's Place in Psychology: LEWIS WAY. (Allen and Unwin, pp. 334, 18s.)

Review of Education in Australia, 1940-48: WADDINGTON, RADFORD and KEATS. (Melbourne Univ. Press, pp. 258, 25s.)

Seven Plus Assessment: C. M. LAMBERT: (i) Reading; (ii) Spelling; (iii) Arithmetic. (Oxford Univ. Press, 25s. per 100+P.T. Manual of Instructions, 1s.)

The Science of Humanity: K. G. COLLIER. (Nelson, pp. 339, 12s. 6d.)

The Educational Thought and Influence of Matthew Arnold: W. F. CONNELL. (Routledge and Kegan Paul, pp. 298, 21s.)

Sixth-Form Citizens (author not stated). (Oxford Univ. Press, pp. 287, 10s. 6d.)

English—The Natural Method: B. L. VULLIAMY. (East Anglian Magazine, Ltd., Vol. I, pp. 114, Vol. II, pp. 153, no price given.)

Frederick Denison Maurice: H. G. WOOD. (Cambridge Univ. Press, pp. 171, 10s. 6d.)

Analytical Psychology and the English Mind: H. G. BAYNES. (Methuen, pp. 242, 18s.)

Colour and Colour Therapy: FABER BIRREN. (McGraw-Hill, 284, 42s. 6d.)

Freud—Dictionary of Psychoanalysis: Edited by NANDOR FODOR and FRANK GAYNOR. (Philosophical Library, New York, pp. 208, 35s. 6d.)

The School Child and his Parents: ZOË BENJAMIN. (Angus and Robertson, pp. 185, 7s. 6d.)

Experimental Design in Psychological Research: ALLEN L. EDWARDS. (Rinehart and Co., Inc., pp. 446, 47s. 6d.)

Selected Readings in Social Psychology: Edited by S. H. BRITT. (Rinehart and Co., Inc., pp. 507, 19s.)

Handbook of Applied Psychology: Edited by D. H. FRYER and E. R. HENRY. (Rinehart and Co., Inc., Vol. 1, pp. 380, Vol. II, pp. 380-842, £5 18s. 0d.)

The Rhesus Danger: R. N. C. McCURDY. (Heinemann, pp. 138, 5s.)

Teaching English through Self-Expression: E. J. BURTON. (Evans Bros., Ltd., pp. 232, 6s.)

Science News: Edited by J. L. CRAMMER. (Penguin, pp. 173, 1s. 6d.)

Economics of Everyday Life: GERTRUDE WILLIAMS. (Pelican, pp. 248, 1s. 6d.)

New Frontiers of the Mind: J. B. RHINE. (Pelican, pp. 220, 1s. 6d.)

National Society for the Study of Education—Forty-ninth Year Book: Part I—Learning and Instruction; Part II—The Education of Exceptional Children. (Cambridge Univ. Press, Pt. I, pp. 352, Pt. II, pp. 346, 22s. 6d. each part.)

Journal d'une Schizophréne: M. A. SECHEHAYE. (Presses Univ. de Fr., pp. 138, 300 fr.)

De la Pratique a la Théorie Psychanalytique: SACHA NACHT. (Pr. Univ. de Fr., pp. 165, 360 fr.)

La Prolongation de la Scholarité: C. MULLER. (Pr. Univ. de Fr., pp. 62, no price given.)

La Logique de l'Assertion: JEAN DE LA HARPE. (Pr. Univ. de Fr., pp. 76, 200 fr.)

Caractérologie des Enfants et des Adolescents: ANDRÉ LE GALL. (Pr. Univ. de France, pp. 458, 800 fr.)

La Pensée Causale en Médicine: WALTHER RIESE. (Pr. Univ. de Fr., pp. 92, 240 fr.)

L'Activité Philosophique Contemporaine en France et aux États-Unis: MARVIN FARBER. (Pr. Univ. de Fr., Part I, pp. 485, 900 fr.; Part II, pp. 408, 800 fr.)

La Société Internationale: THÉODORE RUYSSEN. (Pr. Univ. de Fr., pp. 240, 440 fr.)

Introduction a l'Épistémologie Génetique: JEAN PIAGET. (Pr. Univ. de Fr., pp. 344, 600 fr.)

Index to All Articles and Their Authors in the British Journal of Educational Psychology

Volumes XI to XX

1941—1950

PREPARED BY DORA G. SCHATZ.

I—Subjects. II—Authors

I.—SUBJECTS.

	Vol.	*Part*	*Page*
Abilities, The Future Measurements of: E. L. Thorndike........	XVIII	I	21
Ability and Achievement, The Effects of Pupil Migrancy on, in Rural and Burgh Schools: A. C. Patterson	XX	III	186
Ability and Income: Cyril Burt	XIII	II	83
Adolescence and Some Problems of Youth Training: C. W. Valentine	XIII	II	57
Adolescent, The Education of the Young: The Psychological Implications of the Norwood Report: Cyril Burt	XIII	III	126
Adolescent in the Factory, The: Cora Tenen	XVII	II	72
Adolescents, The Decay of Educational Attainments Among, After Leaving School	XIV	I	19
Adult Studies in Voluntary Courses, The Origins of Interest and Motives for Study of Natural Sciences and Psychology Among: W. E. Flood and R. W. Crossland	XVIII	II	105
Air-Raids, The Effects of, on School Children: Agnes G. McClure..	XIII	I	24
Arithmetic, Problem Solving in, an Investigation into some Aspects of: John Sutherland. Part I	XI	III	215
Part II	XII	I	35
Boredom: Hilde Lewinsky	XIII	III	147
Central School Pupils, The Attitude of, to Certain School Subjects, and the Correlation between Attitude and Attainment: David Jordan	XI	I	29
Child-Adult Relationship, Normalisation of the: Claud A. Claremont	XIV	I	35
Cinema Attendance, The Effects of, on the Behaviour of Adolescents, as seen by their Contemporaries: W. D. Wall and W. A. Simson	XIX	I	53
Cinema Commission of Enquiry Experiment, A Discussion of Results and Experimental Methods: S. J. F. Philpott	XVI	I	32
Clumsiness, What is?: T. H. Pear	XI	II	99
Co-education, Attitudes to Teachers and Scholars Towards: Frank E. Moreton	XVI	II	82
Deaf Children, A Preliminary Survey of some Mental Abilities of: H. Blair-Hood	XIX	III	210
Educational Psychology, American and British—Some Points of Comparison: R. L. Archer	XI	II	128
Educational Psychology and Teaching Method, A List of Researches in: A. M. Blackwell. Part I..........................	XIII	III	153
Part II..........................	XIV	I	46
Part III..........................	XIV	II	99
Part IV..........................	XIV	III	162
Part V..........................	XV	II	93
Educational Research, The Development of, in Great Britain: F. J. Schonell. Part I..........................	XVII	III	131
Part II..........................	XVIII	I	1
Part III..........................	XVIII	II	53
Part IV..........................	XIX	I	16
Part V..........................	XIX	II	82
Part VI..........................	XIX	III	160
Emotional Symptoms in School Children: Jean D. Cummings. Part I: The Incidence of	XIV	III	151
Part II: A Follow-Up Study of	XVI	III	163

	Vol.	Part	Page
English Compositions, A New Method of Marking : R. L. Morrison and P. E. Vernon	XI	II	109
Evacuated Central School Children, Some Sex Differences in Attitudes towards Change of Environment Among : Gertrude Keir	XV	III	146
Evacuated Children, The Billeting of : Cyril Burt	XI	II	85
Evacuation and Air-Raids, A Study of Effects of, on Children of Pre-School Age : Enid John	XI	III	173
Evacuation, Effects of, on the Children : William Boyd	XI	II	120
Evacuation Investigations, Editorial Note on : C. W. Valentine	XI	II	127
Examination, The Position and Size of the Border-line Group in an : A. E. G. Pilliner	XX	II	133
Examinations, An Investigation of the Value of Higher School Certificate Results in Predicting Performance in the First-year University : E. M. Williams	XX	II	83
Farming Camp School, An Experimental : W. S. Flack	XV	I	41
Film Choices of Adolescents, The : W. D. Wall and E. M. Smith	XIX	II	121
Film and Film-Strip in the Instruction of Adults, An Experiment on the Value of : P. E. Vernon	XVI	III	149
Films, The Emotional Responses of Adolescent Groups to Certain : W. D. Wall and W. A. Simson. Part I	XX	III	153
General Information Test, A New : E. Anstey, R. F. Dowse and M. Duguid	XVIII	III	156
Grammar, The Difficulty of : W. J. Macaulay	XVII	III	153
Group Test I.Q's., The Inconsistency of : Hugh McRae	XII	I	59
Hamley, Professor H. R. : Obituary Notice : F. Clarke	XIX	III	151
Handwork Teaching, An Enquiry into : Hector Lamb.			
Part I : *The Relative Popularity of Technical and Expressional Methods of*	XII	I	26
Part II : Their Effects on Character Development	XII	II	108
Havelock Ellis (1859-1939) : J. C. Flugel	XI	I	1
Human Nature, Some Surprising Beliefs Concerning, Among Pre-Medical Psychological Students : Lynn L. Ralya	XV	II	70
Illiterate Adults, The Education of : Cyril Burt	XV	I	20
Industrial Workers, The Educational Interests of a Group of Young : W. D. Wall	XV	III	127
Insecurity and Social Maladjustment in Children : C. Sanders	XVIII	III	148
Instincts, Is the Doctrine of, Dead ? A Symposium :			
Part I : The Case for Human Instincts : Cyril Burt	XI	III	155
Part II : Some Objections to the Theory of Human Instincts : P. E. Vernon	XII	I	1
Part III : Human Instincts and Doctrines about them : E. L. Thorndike	XII	II	85
Part IV : Instinct as Impulse : James Drever	XII	II	88
Part V : Not Dead, but Obsolescent : T. H. Pear	XII	III	139
Part VI : Restropect and Prospect : Charles H. Myers	XII	III	148
Part VII : Conclusion : Cyril Burt	XIII	I	1
Intelligence and Cultural-Educational Level, Differences in, between Different Educational Groups : E. Anstey, R. F. Dowse and M. Duguid	XX	II	99
Intelligence Among Elementary School Children in Northern Ireland, The Distribution of : J. K. Forbes	XV	III	139
Intelligence, Measuring Sudanese : G. C. Scott	XX	I	43
Intelligence Among University and College Students, The Distribution of : Godfrey H. Thomson	XV	II	76
Intelligence in U.P., India, Distribution of : Sohan Lall	XIV	II	95
Intelligence and Related Topics, An Enquiry as to Popular Views on : J. C. Flugel	XVII	III	140
Intelligence Test in an Infant School, The Application of a New, and the Prediction of Backwardness : B. B. Wakelam	XIV	III	142
Intelligence Test Results, The Effect of Practice in : Douglas M. McIntosh	XIV	I	44
Intelligence Test Results, Disclosing Results of to Non-Professional Persons : P. E. Vernon	XIV	II	106
Interest in Practical Topics, Assessment of : E. A. Peel	XVIII	I	41
Interests of Senior School Children in Non-Scientific Subjects : R. Rallison	XIII	I	39

	Vol.	Part	Page
Juvenile Delinquency, Recent Discussions of: Cyril Burt	XIX	I	32
Languages, A Comparison of the Orders of Merit of Higher School Certificate Candidates Offering Two Modern: James A. Petch	XV	III	133
Latin Square, The Use of, in Designing Educational Experiments: Godfrey H. Thomson	XI	II	135
Leisure Activities of Grammar School Children, The: Mary Stewart	XX	I	11
Maladjusted Children of High Intelligence: C. L. C. Burns	XIX	II	137
Marking of English Composition in Grammar School Selection, The: S. Wiseman	XIX	III	200
Matrices, An Experimental Survey with Progressive: G. A. Foulds and J. C. Raven	XX	II	104
Mental Abilities and Mental Factors: Cyril Burt	XIV	II	85
Mental Quality of the Population, Psychological Studies of the: P. E. Vernon	XX	I	35
Musical Ability, The Assessment of: James Mainwaring	XVII	II	83
Musicianship, The Meaning of: James Mainwaring	XI	III	205
Myers, Charles C.: T. H. Pear	XVII	I	1
Nazi School System, The Social and Educational Background of the: Hans Liebeschuetz	XI	III	197
Newspaper Reading of Adolescents and Adults: W. D. Wall.			
Part I	XVIII	I	26
Part II	XVIII	II	87
Nunn, Sir Percy: H. R. Hamley	XV	I	1
Parent-Teacher Co-operation, The Opinions of Teachers on: W. D. Wall	XVII	II	97
Parents, The Psychological Guidance of, in the Upbringing of Children: A Symposium and Discussion: C. W. Valentine and Grace Rawlings	XVI	II	96
Personality—A Symposium:			
Part I: The Assessment of Personality: Cyril Burt	XV	III	107
Part II: Personality of the Problem Child: Alan Maberley	XVI	I	5
Part III: Geneticism versus Ego-Structure in Theories of Personality: Gordon W. Allport	XVI	II	57
Part IV: Both Sides of the Shield; The Reactions of an Outsider: Godfrey H. Thomson	XVI	III	105
Part V: Reply to Criticisms, and Conclusions: Cyril Burt	XVII	I	6
Physical Training on Personality, An Experimental Investigation into the Effects of: F. J. Milverton	XIII	I	30
Physics, Objective Test Forms for School Certificate: H. P. Wood	XIII	III	141
Poetry Discrimination, A New Test of: E. M. Eppel	XX	II	111
Probationers, An Analysis of the Records of Some 750: E. W. Hughes	XIII	III	113
Psychology, Post Graduate Training of Teachers in: P. E. Vernon	XX	III	149
Psychology, Some Present-Day Trends, Dangers and Possibilities in the Field of—Presidential Address to the British Psychological Society: C. W. Valentine	XVIII	III	134
Readers, Teaching Backward: Cyril Burt and R. B. Lewis	XVI	III	116
Reading Ability, The Improvement in: L. B. Birch	XX	II	73
Reading Ability and I.Q. as Measured by a Verbal Group Intelligence, An Investigation into the Relationship between: Margaret A. Mellone	XII	II	128
Reading and Arithmetic Tests for Seven-year-Olds, New Norms for Ballard's: Godfrey H. Thomson and D. N. Lawley	XII	I	56
Reading Backwardness among Men in the Army: W. D. Wall.			
Part I	XV	I	28
Part II	XVI	III	133
Reading, Comprehension in Silent: Harry Bell	XII	I	47
Reading Disability, The Relation of, to Handedness and Certain Ocular Factors: F. J. Schonell	XI	I	20
Religion Among Adolescents and Adults, Attitudes to: F. E. Moreton	XIV	II	69
Research in the New Setting, Educational: F. Clarke	XIV	I	1
School Attendance, The Relation of, to Employment Records, Army Conduct and Performance in Tests: Clifford Lummis	XVI	I	13
Secondary Education, Psychological and Sociological Principles for the Re-organization of: A. Pinsent	XIV	II	57

	Vol.	Part	Page
Selection of Pupils for Different Types of Secondary Schools; Symposium on the:			
Part I: A General Survey: Cyril Burt	XVII	II	57
Part II: An Administrator's Point of View: W. P. Alexander	XVII	III	123
Part III: A Method of Allocation in a County Borough: V. J. Moore	XVIII	I	16
Part IV: An Experimenter's Point of View: E. J. G. Bradford	XVIII	II	67
Part V: The Selector's Point of View: J. J. Dempster	XVIII	III	121
Part VI: Evidence of a Practical Factor at the Age of Eleven: E. A. Peel	XIX	I	1
Part VII: A Survey of Ability and Interest at the Stage of Transfer: C. M. Lambert	XIX	II	67
Part VIII: An Industrial Psychologist's Point of View: A. Rodger	XIX	III	154
Part IX: Conclusion: Cyril Burt	XX	I	1
Social Differences in English Education: T. H. Pear	XIV	III	113
Socio-Economic Level and Test Performance: C. M. Fleming	XIII	II	74
Spearman, Professor C.: P. B. Ballard	XVI	I	1
Stour, George Frederick (1860-1944), An Appreciation: Rex Knight	XVI	II	53
Structure of the Mind, The: A Review of the Results of Factor Analysis: Cyril Burt. Part I	XIX	II	100
Part II	XIX	III	176
Student Selection by Means of Psychological Tests—A Critical Survey: H. J. Eysenck	XVII	I	20
Sublimation, its Nature and Conditions: J. C. Flugel. Part I	XII	I	10
Part II	XII	II	97
Part III	XII	III	162
Teachers, The Social and Educational Background of Emergency Trained, and Reasons for their Choice of the Profession: P. J. T. Grant	XX	III	164
Teachers' Assessment of their Pupils, The Reliability of: Cyril Burt	XV	II	80
Teachers' Diploma Examination, The Relative Performances of Arts and Science Graduates in a: H. P. Wood	XI	I	8
Teaching, Fallacious Arguments from Experiments on Methods of: S. F. J. Philpott	XV	II	57
Teaching Mark as a Criterion of Future Success in the Teaching Profession; A Study of the Training College Final: William B. Tudhope. Part I	XII	III	167
Part II	XIII	I	16
Teaching Profession, Motives for the Choice of, by Training College Students: William B. Tudhope	XIV	III	129
Teaching Success of University Students, A Further Investigation of Pre-College Teaching Experience and other Factors in the: Hywela Saer	XI	III	183
Teaching and Teacher Training, The Psychology of: A Review in the Relation to the McNair Proposals: A. Pinsent	XV	I	5
Technical Education, Selection for: E. J. G. Bradford. Part I	XVI	I	20
Part II	XVI	II	69
Terman Binet Tests, A Factorial Analysis of: Cyril Burt and Enid John. Part I	XII	II	117
Part II	XII	III	156
Training College Course, Attitudes to Secondary School Authorities Towards, Encountered by Intending Students: William B. Tudhope	XIV	I	7
Visual Presentation of Factual Data, The: M. D. Vernon	XX	III	174
Vocabulary, The Estimation of Difficulty of: P. E. Vernon	XX	II	77
Welsh as a Second Language, Attitude Towards: A Preliminary Investigation: W. R. Jones	XIX	I	44
Welsh as a Second Language: W. R. Jones	XX	II	117
Young People's Colleges and Leisure Time: Observer	XIV	II	80
Youth Leader Selection, Group Methods applied to: A. G. Beverstock	XIX	II	112
Youth, The Service of: Olive A. Wheeler	XIII	II	69
Youth Centre Interests, An Analysis of: W. Hammond	XV	III	122

II.—INDEX OF AUTHORS (VOLUME AND PAGE).

Alexander, W. P.: XVII, 57 and 123.
Allport, Gordon W.: XVI, 57.
Anstey, E., Dowse, R. F., and Duguid, M.: XVIII, 156; XX, 99.
Archer, R. L.: XI, 128.
Ballard, P. B.: XVI, 1.
Bell, Harry: XII, 47.
Beverstock, A. G.: XIX, 112.
Blackwell, A. M.: XIII, 153; XIV, 46, 99, and 162; XV, 93.
Blair-Hood, H.: XIX, 210.
Bradford, E. J. G.: XVI, 20 and 69; XVIII, 67.
Burns, C. L. C.: XIX, 137.
Burt, Cyril: XIII, 1, 83, 126; XIV, 85; XV, 20, 80, and 107; XVII, 6, 57; XIX, 32, 100, and 176; XX, 192.
Burt, Cyril and Lewis, B.: XVI, 116.
Burt, Cyril and John, Enid: XII, 117 and 156.
Boyd, William: XI, 120.
Claremont, Claude A.: XIV, 35.
Clarke, Fred: XIV, 1; XIX, 151.
Crossland, R. W., and Flood, W. E.: XVIII, 105.
Cummings, Jean D.: XIV, 151; XVI, 163.
Dempster, J. J.: XVIII, 121.
Dowse, R. F., Anstey, E., and Duguid, M.: XVIII, 156; XX, 99.
Duguid, M., Anstey, E., and Dowse, R. F.: XVIII, 156; XX, 99.
Drever, James: XII, 88.
Eppel, E. M.: XX, 111.
Eysenck, H. J.: XVII, 20.
Flack, W. S.: XV, 41.
Fleming, C. M.: XIII, 74.
Flood, W. E., and Crossland, R. W.: XVIII, 105.
Flugel, F. C.: XII, 10, 97, and 162; XVII, 140.
Forbes, J. K.: XV, 139.
Foulds, G. A., and Raven, J. C.: XX, 104.
Grant, P. J. T.: XX, 164.
Hamley, H. R.: XV, 1.
Hammond, W.: XV, 122.
Hughes, E. W.: XIII, 113.
John, Enid: XI, 173.
John, Enid, and Burt, Cyril: XII, 117 and 156.
Jones, W. R.: XIX, 44; XX, 117.
Jordan, David: XI, 28.
Keir, Gertrude: XV, 146.
Knight, Rex: XVI, 53.
Lall, Sohan: XIV, 95.
Lamb, Hector: XII, 26 and 108.
Lambert, C. M.: XIX, 67.
Lawley, D. N. and Thomson, Godfrey H.: XII, 56.
Lewinsky, Hilde: XIII, 147.
Lewis, R. B., and Burt, Cyril: XVI, 116.
Liebeschuetz, Hans: XI, 197.
Lummis, Clifford: XVI, 13.
Maberly, Alan: XVI, 5.
Macaulay, W. J.: XVII, 153.
Mace, C. A.: XVI, 39.
Mainwaring, James: XVII, 83.
Mellone, Margaret A.: XII, 128.
Milverton, F. J.: XIII, 30.
Moore, V. J.: XVIII, 16.
Moreton, F. E.: XIV, 69: XVI, 82.
Morrison, R. L., and Vernon, P. E.: XI, 109.
Myers, Charles H.: XII, 148.
McClure, Agnes G.: XIII, 24.
McIntosh, Douglas M.: XIV, 44.
McRae, Hugh: XII, 59.
Observer: XIV, 80.
Patterson, A. C.: XX, 186.
Pear, T. H.: XIV, 113; XVII, 1.
Peel, E. A.: XVIII, 41; XIX, 1.
Petch, James A.: XV, 133.
Philpott, S. J. F.: XV, 57; XVI, 32.
Pilliner, A. E. G.: XX, 133.
Pinsent, A.: XIV, 57; XV, 5.
Rallison, R.: XIII, 39.
Ralya, Lynn L.: XV, 70.
Raven, J. C., and Foulds, G. A.: XX, 104.
Rawlings, Grace and Valentine, C. W.: XVI, 96.
Rodger, A.: XIX, 154.
Saer, Hywela: XI, 183.
Sanders, C.: XVIII, 148.
Schonell, F. J.: XI, 20; XVII, 131; XVIII, 1 and 53; XIX, 16, 82, and 160.
Scott, G. C.: XX, 43.
Simson, W. A., and Wall, W. D.: XIX, 53; XX, 153.
Smith, E. M., and Wall, W. D.: XIX, 121.
Stewart, Mary: XX, 11.
Sutherland, John: XI, 215; XII, 35.
Tenen, Cora: XVII, 72.
Thomson, Godfrey H.: XI, 135; XV, 76; XVI, 105.
Thomson, Godfrey H. and Lawley, D. N.: XII, 56.
Thorndike, E. L.: XII, 85; XVIII, 21.
Tudhope, William B.: XII, 167; XIII, 16; XIV, 129.
Valentine, C. W.: XI, 127; XVIII, 134.
Valentine, C. W., and Rawlings, Grace: XVI, 96.
Vernon, M. D.: XX, 174.
Vernon, P. E.: XII, 1; XIV, 106; XVI, 149; XX, 35 and 149.
Wakelam, B. B.: XIV, 142.
Wall, W. D.: XIV, 19; XV, 28 and 127; XVI, 133; XVII, 114; XVIII, 28 and 86.
Wall, W. D., and Simson, W. A.: XIX, 53; XX, 153.
Wall, W. D., and Smith, E. M.: XIX, 121.
Wheeler, Olive A.: XIII, 69.
Williams, E. M.: XX, 83.
Wiseman, S.: XIX, 200.
Wood, H. P.: XI, 8; XIII, 141.